

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2001, 17:33:32 ; Search time 4677.26 Seconds

(without alignments)
13261.806 Million cell updates/sec

Title: US-09-041-994-1

Perfect score: 4496

Sequence: 1 GCTGAGTGTGCTGACTCAGAC.....CATTTGACGACGATTTCTAG 4496

Scoring table: IDENTITY_NUC

Gap 10.0 , Gapext 1.0

Searched: 13168883 seqs, 6898233199 residues

Total number of hits satisfying chosen parameters: 26337766

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/US0612_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US096_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US098_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US099_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099G_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099H_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099I_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4496	100.0	4496	US-09-041-994-1	Sequence 1, Appl1
2	4494.4	100.0	4496	US-09-513-066-1	Sequence 1, Appl1
3	4451.6	99.0	6835	US-09-125-635-1	Sequence 1, Appl1
4	4434.2	98.6	4789	US-60-068-511-1	Sequence 1, Appl1
5	4429.6	98.5	6855	US-09-359-992-3962	Sequence 3962, Ap
6	4429.6	98.5	6855	US-09-359-992-3962	Sequence 3962, Ap
7	4421.2	98.3	7956	US-60-172-373-13551	Sequence 13551, A
8	4393.8	97.7	7116	US-60-209-009-198	Sequence 198, App
9	4393.8	97.7	7116	US-60-213-360-7991	Sequence 7991, Ap
10	4322.2	96.1	6754	US-09-440-612-1	Sequence 1, Appl1
11	4322.2	96.1	6754	US-09-440-612-1	Sequence 1, Appl1
12	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
13	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
14	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
15	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
16	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
17	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
18	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
19	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
20	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
21	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
22	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
23	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
24	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
25	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
26	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
27	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
28	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
29	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
30	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
31	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
32	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
33	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
34	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
35	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
36	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
37	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
38	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
39	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
40	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
41	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
42	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
43	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
44	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap
45	4313	95.9	7888	US-09-652-123-8947	Sequence 8947, Ap

ALIGNMENTS

RESULT 1
US-09-041-994-1
Sequence 1, Application US/09041994
GENERAL INFORMATION:
APPLICANT: Chen, J. Don
APPLICANT: Li, Hui
TITLE OF INVENTION: Transcriptional Coactivator for Nuclear
TITLE OF INVENTION: Hormone Receptors
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lohive and Cockfield
STREET: 28 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,994
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Liepmann, W. Hugo
REGISTRATION NUMBER: 20,407
REFERENCE/DOCKET NUMBER: UMN-026-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-227-7400
TELEFAX: 617-742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4496 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 86..4338
US-09-041-994-1
Query Match 100.0%; Score 4496; DB 14; Length 4496;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 301 AAGGAACAGTAAGACAGATACGTCAATAATAAAGCAGAGAAAGCAACTATTTCACATGA 360
OY 361 TGTATGATGTTCAAAAAGCCGATGTATCTTCACAGGGCAGGGAGTTATTGATTAAGACTC 420
Db 361 TGTATGATGTTCAAAAAGCCGATGTATCTTCACAGGGCAGGGAGTTATTGATTAAGACTC 420
OY 421 CTYAGACCGCTTTTACCTTCAGGCAATGAGGTTGCTTATTTGTGTGAATCGAGAGGC 480
Db 421 CTYAGACCGCTTTTACCTTCAGGCAATGAGGTTGCTTATTTGTGTGAATCGAGAGGC 480
OY 481 AATCAATGATTTTGTATCAGAAAATGTCACACAAATACCTGCATATTAAGCAAGAGACT 540
Db 481 AATCAATGATTTTGTATCAGAAAATGTCACACAAATACCTGCATATTAAGCAAGAGACT 540
OY 541 GGTAAACACAAAGTGTTCACATATCTTCATGAAAGAGAGAGATTTTCTTAAGAA 600
Db 541 GGTAAACACAAAGTGTTCACATATCTTCATGAAAGAGAGAGATTTTCTTAAGAA 600
OY 601 TTTACCAAAATCTACAGTTAATGAGTTTCTGACAAATGAGCCCAAGCAAAAAAG 660
Db 601 TTTACCAAAATCTACAGTTAATGAGTTTCTGACAAATGAGCCCAAGCAAAAAAG 660
OY 661 CCATACATTTTATTTGCGCGTATGATGAAAAACACACATGATATTTCGGAAGACATAA 720
Db 661 CCATACATTTTATTTGCGCGTATGATGAAAAACACACATGATATTTCGGAAGACATAA 720
OY 721 CGCCAGTCCTGAAATGCGCCAGAGATATGAAACAATGCAAGTCTTGCCCTGTCCAGCC 780
Db 721 CGCCAGTCCTGAAATGCGCCAGAGATATGAAACAATGCAAGTCTTGCCCTGTCCAGCC 780
OY 781 ACGACCTATGATGAGGAAGGGGAGATTGCAATCTTGATGATCTGTGGCAGCGCG 840
Db 781 ACGACCTATGATGAGGAAGGGGAGATTGCAATCTTGATGATCTGTGGCAGCGCG 840
OY 841 CATTACTACAGAGAAAAGCAATTTTCCATCAACCTTGAGAGCTTTATTACCAACATGA 900
Db 841 CATTACTACAGAGAAAAGCAATTTTCCATCAACCTTGAGAGCTTTATTACCAACATGA 900
OY 901 TCTTTACAGAAAAGTGTTCATATATAGATACAAATTCAGTGAAGTCTCATGAGCGCTGG 960
Db 901 TCTTTACAGAAAAGTGTTCATATATAGATACAAATTCAGTGAAGTCTCATGAGCGCTGG 960
OY 961 CTTTAAGATATATATCCGAAGGTGATTCAGAGATTTTATAGTCTAATATGATGGGAGTC 1020
Db 961 CTTTAAGATATATATCCGAAGGTGATTCAGAGATTTTATAGTCTAATATGATGGGAGTC 1020
OY 1021 ATGGTCCAGAAAGCTCACTATCAAGAAAGCTTATCTTAATGGCCATGCAAGAAACCCAGT 1080
Db 1021 ATGGTCCAGAAAGCTCACTATCAAGAAAGCTTATCTTAATGGCCATGCAAGAAACCCAGT 1080
OY 1081 ATATGATTTCTGTTGGCTGATGGAAGTATAGTGAAGTCTGACAGACAAACCAACTCTT 1140
Db 1081 ATATGATTTCTGTTGGCTGATGGAAGTATAGTGAAGTCTGACAGACAAACCAACTCTT 1140
OY 1141 CCGAAATCTCTGTACAAATGATGACATGCGTTGTCTCAACCCACTCTTTCAGAGAGA 1200
Db 1141 CCGAAATCTCTGTACAAATGATGACATGCGTTGTCTCAACCCACTCTTTCAGAGAGA 1200
OY 1201 ACAGATGATATATACCAAAACCCAAATCTGTTGGAGAAAGGATTAGACCACTATGGC 1260
Db 1201 ACAGATGATATATACCAAAACCCAAATCTGTTGGAGAAAGGATTAGACCACTATGGC 1260
OY 1261 TGGATGCAACAGTTTCGTAGAGCGGATGAGTATGTCGCAACCAAGCTTACAGATGCC 1320
Db 1261 TGGATGCAACAGTTTCGTAGAGCGGATGAGTATGTCGCAACCAAGCTTACAGATGCC 1320
OY 1321 GAGCAGACAGGGCTATGCGCTTGGCAGACCTTAGACACACAGGGCAGATGAGTGAAGTAG 1380
Db 1321 GAGCAGACAGGGCTATGCGCTTGGCAGACCTTAGACACACAGGGCAGATGAGTGAAGTAG 1380
OY 1381 GTATGGGGGTTCCAGTAAACATAGCTCATTTGACCCCTGGGCCAGGCAATGCAATCACCTC 1440
Db 1381 GTATGGGGGTTCCAGTAAACATAGCTCATTTGACCCCTGGGCCAGGCAATGCAATCACCTC 1440

Db 1381 GTATGGGGGTTCCAGTAACTACCTTCAATGACCCCTGGGCCAGCATGCATCACCATC 1440
QY 1441 TTCTTACGAAACACACATGAGGCTCAACATGAGTACGCCCCCATGAGGCTCTGG 1500
Db 1441 TTCTTACGAAACACACATGAGGCTCAACATGAGTACGCCCCCATGAGGCTCTGG 1500
QY 1501 TCTTGGCCCAACACAGAGATATCATGATTTCTCTCGTAATCGTGGAGTCCAAAGAT 1560
Db 1501 TCTTGGCCCAACACAGAGATATCATGATTTCTCTCGTAATCGTGGAGTCCAAAGAT 1560
QY 1561 AGCCTACATCACTTTTCTCTCTGTCAGAGTGTGACCTCTCCATGAGCATCTTGGCAA 1620
Db 1561 AGCCTACATCACTTTTCTCTCTGTCAGAGTGTGACCTCTCCATGAGCATCTTGGCAA 1620
QY 1621 TACTGGGAACACAGCTTTTCCAGACGCTCTCAGTGTGAGGCTCCAAAGCCATGTAAGG 1680
Db 1621 TACTGGGAACACAGCTTTTCCAGACGCTCTCAGTGTGAGGCTCCAAAGCCATGTAAGG 1680
QY 1681 TGTGGGGACTTCCCTTTTATCTACTCTGTCATCAGGAGGCCCCAAATGGATTAATCTCC 1740
Db 1681 TGTGGGGACTTCCCTTTTATCTACTCTGTCATCAGGAGGCCCCAAATGGATTAATCTCC 1740
QY 1741 CAATATGAATATTAATCCCAACCAAGTAAGTAAGCAATCAGATTCGAAGTCTCTGGG 1800
Db 1741 CAATATGAATATTAATCCCAACCAAGTAAGTAAGCAATCAGATTCGAAGTCTCTGGG 1800
QY 1801 CTTTATTTGCGACCAAAATTCAGTGTGAGAGTAAATGTGTCAGTCAATATGAGAGATCA 1860
Db 1801 CTTTATTTGCGACCAAAATTCAGTGTGAGAGTAAATGTGTCAGTCAATATGAGAGATCA 1860
QY 1861 CCTCAGTACAAAGAAAGTAAGAGAGAGAGTGTGAGGGGGCAGAGATCAAGGGGGTCC 1920
Db 1861 CCTCAGTACAAAGAAAGTAAGAGAGAGAGTGTGAGGGGGCAGAGATCAAGGGGGTCC 1920
QY 1921 TTTGGAAGCAAGAGTCAATAAAAATTAATCTGAGTACTTACCTGTTCTCTGATGACG 1980
Db 1921 TTTGGAAGCAAGAGTCAATAAAAATTAATCTGAGTACTTACCTGTTCTCTGATGACG 1980
QY 1981 GGGTCACTTCTCTCTTACCAACCTCCCTGAGTTCAGTGTGTAAGAAATCTTCTGTAG 2040
Db 1981 GGGTCACTTCTCTCTTACCAACCTCCCTGAGTTCAGTGTGTAAGAAATCTTCTGTAG 2040
QY 2041 TGTCAACAGCCCTCTGAGAGTCTCTCTCTCTACATCTGAGAGGATTCCTCTCATATCAA 2100
Db 2041 TGTCAACAGCCCTCTGAGAGTCTCTCTCTCTCTACATCTGAGAGGATTCCTCTCATATCAA 2100
QY 2101 TATGATGGGTCTACTGTACAGAGAACGACCGATTTTGCACAAAGTGTCTGACAGATG 2160
Db 2101 TATGATGGGTCTACTGTACAGAGAACGACCGATTTTGCACAAAGTGTCTGACAGATG 2160
QY 2161 GAATTCACAGCTGAGGTAGCCAAAGTACTGTGACAAAGCCACTGGGAAAGACACAGCAG 2220
Db 2161 GAATTCACAGCTGAGGTAGCCAAAGTACTGTGACAAAGCCACTGGGAAAGACACAGCAG 2220
QY 2221 TATTAATCTTGTGGGAGCGAAATGTGTCAAGAGAGAGACGTAAGTCTTAAGAAAG 2280
Db 2221 TATTAATCTTGTGGGAGCGAAATGTGTCAAGAGAGAGACGTAAGTCTTAAGAAAG 2280
QY 2281 GGAATTAATGACCTTTTATGATACCTGCTGACAGAGGATGATCTAGTATGACACTCTC 2340
Db 2281 GGAATTAATGACCTTTTATGATACCTGCTGACAGAGGATGATCTAGTATGACACTCTC 2340
QY 2341 TTAAGAACTAACGCCCCAGTGAAGAGAGTGAACATTAATATGATGACAGCAGCTC 2400
Db 2341 TTAAGAACTAACGCCCCAGTGAAGAGAGTGAACATTAATATGATGACAGCAGCTC 2400
QY 2401 CACCATTTCTAGCTCAAGTCAAGAGAAAGACCTTAATATTAAGACAGAGCAAGTAAAG 2460
Db 2401 CACCATTTCTAGCTCAAGTCAAGAGAAAGACCTTAATATTAAGACAGAGCAAGTAAAG 2460
QY 2461 GGGATCTGAGAGCTTGATTAATCTAGATGCTATTTGGTGTGATCTAGTCTGACTT 2520
Db 2461 GGGATCTGAGAGCTTGATTAATCTAGATGCTATTTGGTGTGATCTAGTCTGACTT 2520

QY 2521 TTCAATTAATTCATATCCCAATGATAGTATCTATGAGGAGTAAACAGAGGTTTCA 2580
Db 2521 TTCAATTAATTCATATCCCAATGATAGTATCTATGAGGAGTAAACAGAGGTTTCA 2580
QY 2581 AGGAATTAATTCCTGGGTTTGAAGATTCACAGTGTGAGTCAATTCCTCTCATTA 2640
Db 2581 AGGAATTAATTCCTGGGTTTGAAGATTCACAGTGTGAGTCAATTCCTCTCATTA 2640
QY 2641 TAAACGAGAGTGTCTGTGATAGCCCTGTTCTGTGGCTCAAGTCTTCCAGTAAAG 2700
Db 2641 TAAACGAGAGTGTCTGTGATAGCCCTGTTCTGTGGCTCAAGTCTTCCAGTAAAG 2700
QY 2701 TATCAGTGTCTTCCATGATTAACAAAGCAACCATGTTGGGGGAATCCAGATGAT 2760
Db 2701 TATCAGTGTCTTCCATGATTAACAAAGCAACCATGTTGGGGGAATCCAGATGAT 2760
QY 2761 GATATGAGGAAATTAATGATGATAGTGGGCTCAAGTATGAGGCTCAAGTATGATGAC 2820
Db 2761 GATATGAGGAAATTAATGATGATAGTGGGCTCAAGTATGAGGCTCAAGTATGATGAC 2820
QY 2821 TCAGACTCTTCTCTCAGAGAGTGGGGCTTACCAAACTCAAGGCGGAGAAATGGAAC 2880
Db 2821 TCAGACTCTTCTCTCAGAGAGTGGGGCTTACCAAACTCAAGGCGGAGAAATGGAAC 2880
QY 2881 TATGAATTCAAATCCATGAGGAGAGACAGAGAGATTAATATCTTCTTACCCAGAC 2940
Db 2881 TATGAATTCAAATCCATGAGGAGAGACAGAGAGATTAATATCTTCTTACCCAGAC 2940
QY 2941 TGCAGTGGGGCTCTATTTCCCATATTCCTCTTCCGTTCAATAGATACAGGTGCGAG 3000
Db 2941 TGCAGTGGGGCTCTATTTCCCATATTCCTCTTCCGTTCAATAGATACAGGTGCGAG 3000
QY 3001 ACCAGTATTCACACAGCAGCAGAGATGCTCAATAGAGGCTGTGAATCCCATGGG 3060
Db 3001 ACCAGTATTCACACAGCAGCAGAGATGCTCAATAGAGGCTGTGAATCCCATGGG 3060
QY 3061 AATGGGGGCTAATCCCTATGAGGCAAGCAGCAGAGATTAACCAATCGGTTCTGCGCCCA 3120
Db 3061 AATGGGGGCTAATCCCTATGAGGCAAGCAGCAGAGATTAACCAATCGGTTCTGCGCCCA 3120
QY 3121 TGGCATGTGTCCATGAGAACAAATTTCTCATGAGCACTCAAAATAGGCTCTTATAGAA 3180
Db 3121 TGGCATGTGTCCATGAGAACAAATTTCTCATGAGCACTCAAAATAGGCTCTTATAGAA 3180
QY 3181 TTCCCTGATGATCTTGTGGGCCACTTCCAACTGGAAGGCGCAGATGACGAAAGAG 3240
Db 3181 TTCCCTGATGATCTTGTGGGCCACTTCCAACTGGAAGGCGCAGATGACGAAAGAG 3240
QY 3241 AATATGAGACCTGACACACTTCTCTCAGCAACACAGATGCGACAGGCTGGAAGAAAT 3300
Db 3241 AATATGAGACCTGACACACTTCTCTCAGCAACACAGATGCGACAGGCTGGAAGAAAT 3300
QY 3301 TGACAGAGCTTTGGGATTCCTGAATCTGTCATGAGGAGCAGGATTAAGAGCCAAACA 3360
Db 3301 TGACAGAGCTTTGGGATTCCTGAATCTGTCATGAGGAGCAGGATTAAGAGCCAAACA 3360
QY 3361 GGAATCTTTCAGAGGCCCAAGAACAGCAGATTAATGATGATCAGAAAGCAGATTAATG 3420
Db 3361 GGAATCTTTCAGAGGCCCAAGAACAGCAGATTAATGATGATCAGAAAGCAGATTAATG 3420
QY 3421 GATGCTTTCAGAGGCCCAAGAACAGCAGATTAATGATGATCAGAAAGCAGATTAATG 3480
Db 3421 GATGCTTTCAGAGGCCCAAGAACAGCAGATTAATGATGATCAGAAAGCAGATTAATG 3480
QY 3481 ACCATTTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Db 3481 ACCATTTTAACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
QY 3541 AATGACCCAGAGCCAGACATCATGAGACCCCGGACAAACACCCCAAGCAACTTGAAT 3600
Db 3541 AATGACCCAGAGCCAGACATCATGAGACCCCGGACAAACACCCCAAGCAACTTGAAT 3600

Db 781 acgagctatgatgaggaagggaatttgcaattctgtatgattctgtgtgcaagccg 840
Qy 841 CATTTACTACAGAGAAAGACATTTTCCATCAAAACCTGAGAGCTTTATTAACGACATGA 900
Db 841 cttactaagagaagaacattccattcaaaccttgagagcttatttaccagatatga 900
Qy 901 TCTTTAGGAAAGGTTGTCAATATAGTACAAATTCAGAGATTCCTCCATGAGGCTGG 960
Db 901 tctctcaagaaaggtctgtcaatatagatacaaatcagatgagatccctcagagccctgg 960
Qy 961 CTTTGAAGATATATATCCGAAAGTGTATTCAGAGATTTTATGCTAATATGATGGGAGTC 1020
Db 961 ctttgaagataataatccgaaggtgtatctcaagatttttagtctaataigtatggcaagtc 1020
Qy 1021 ATGCTCCAGAAAGCTCTATATCAAGAGCTTATCTTAAATGGCCATGCAAAACCCAGT 1080
Db 1021 atgtctccagaaagctcactataagaagcttacttaatgtgcatgtcagaaacccagtc 1080
Qy 1081 ATATGATTTCTGTTGGCTGATGAACTATATGATGCTGACAGACAAAAGCAAGCTTT 1140
Db 1081 atatcgattctcgctgtgcgtgagtggaactatagtgactgtgcagacaaagcaaacctctc 1140
Qy 1141 CCGAAATCCTGTAAACAATGATGACATGGCTTTGTCTCAACCACCTCTCTTCAGAGAGA 1200
Db 1141 ccgaaatccctgtacaataigtatgacatgtgcttctgtctcaaccactctctctcagagaga 1200
Qy 1201 ACAGAAATGATATATAGCAAAACCCAAATCTGTTGGACAAGGATTAAGACCACTATGCC 1260
Db 1201 acagaaatgatataagacaaacccaatccctgtgtgacaagggatlagaccactatgtgc 1260
Qy 1261 TGATGCAACAGTTGCTGGTAGGGGCAATGATGATGCTGCAACCAAGGCTTACAGATGCC 1320
Db 1261 tggatgcaacagttgctggtaggggcaatgattgattgctgcaacaaagccttaccagatgcc 1320
Qy 1321 GAGCAGACAGGCTTATGCTGTTGGCAGACCTTACACACAGGCGAGATGATGGAGCTAG 1380
Db 1321 gagcagacagggcttattgctgttggcagaccttgcagaccacagggcagatgagtgagagctag 1380
Qy 1381 GTATGGGGGTTCCAGTATACATAGCTTCATATTGACCCCTGGGCGCAGCATGCAATCACCATC 1440
Db 1381 gtatgggggttccagtataacataagcttcatgtaccctgtggtccaggtcagatgcaatcacatc 1440
Qy 1441 TTCTCTACAGAAACAACATATGGGTCACATGATGATGAGTACGCCCCACATGGAGTCTCG 1500
Db 1441 ttctctacagaaacaacaacataaggtgcacaagaagtaagcccccaatgagggatccctgg 1500
Qy 1501 TCTTGCCCCAAMCAGACAGATATCATGATTTCTCTCTGTAATGTTGGAGATCCAAAGAT 1560
Db 1501 tcttgccccaaacagcagatatcatgtattctctcgttaatcgtgaggtcccaaatg 1560
Qy 1561 AGCCTACATCATGTTTCTCTCTGTTGAGGTGTGACATCTCCATGCAATCTTCTGGCAA 1620
Db 1561 agcctacatcaagttctctcgttgcaggtgtgcactctcccatgtgcatcttcttgcgaa 1620
Qy 1621 TACTGGGAACCAAGCTTTTCCAGAGCTCTCAGTGCCTCTCAAGCCATCAGTGAAG 1680
Db 1621 tactgggaaccaaagcttttccagagctctctcagtgccctgcagaccatcagtgaaag 1680
Qy 1681 TGTGGGACTTCCCTTTTATCTACTGTGTCATCAGCAGGCCCCAAATTTGATTAATCTTCC 1740
Db 1681 tgtgggacttcccttttattactactctgtcatcacaggccccaaattgtgataactcttc 1740
Qy 1741 CAATATGAATATATACCAACCAAGTAAAGTAAACATCAGAGATTTCCAAAGTCTCTGGG 1800
Db 1741 caataatgaatatataccaacaagaagtaagaatcaagatcccaagagctctctcgtg 1800
Qy 1801 CTTTATTTGAGACCAAAATCCAGTGAAGTGAATGCTGCACTCAAAATGACAGAGATCA 1860
Db 1801 cttttatgtgacaaaatccatccagtgagaggttcaatgtgtcagttcaaatagccagagatca 1860
Qy 1861 CCTCAGTGACAAAGAAAGTAAGAGAGACAGTGTGAGGGGGCAGAGAAATCAAGGGGTCC 1920
Db 1861 cctcagtgacaaaagaagtaagagagcagtggtgagggggcagagatccaaaggggtcc 1920
Qy 1921 TTTGGAAGCAAGGTCATTAATAAATTACTGCAAGTTACTTACTGTTCTTCTGATGACCG 1980
Db 1921 ttggaagaagaaagtcataaataattactgcagttacttactactcttctctgtgagcg 1980
Qy 1981 GGGTCATTCCTCTTACCAACCAATCCCCCTAGATTCAGTTCAGTGAAGTCAATCTTGTAG 2040
Db 1981 gggtcattctctcttgcagcaactccccctagatltcaagttgttaagaatcttctgttag 2040
Qy 2041 TGTACACAGCCCTCTGAGATCTCTCTCTATCTGAGAGATATCTTATCATCCAA 2100
Db 2041 tgtcacagccctctgagagcttccctctccctacatctgagaggtatccctctacatccaa 2100
Qy 2101 TATGATGGGTCTGCTGTACAGAGAACGACCGGATTTTGACAAAGTTGCTGCAAGATG 2160
Db 2101 tatgcatgggtctgctgttacagagaaagcagcgatttgcacaaatgtctgcagaaatgg 2160
Qy 2161 GAATTCACACAGCTGAGATGACCAAGATTAATCTGCACAGCCACTGGGAAAGACACGACAG 2220
Db 2161 gaattcacacagctgaggttagccaagattactgtcacaaagccactggtgaaagaacacagcg 2220
Qy 2221 TATTAATCTTGTGTGGGACGGAATGTTGTCAAGACAGAGACAGTAACTTAAGAGAA 2280
Db 2221 tataacttcttgtgggagcggaatgtgtcagaagcagagcagctaaagtccctaagaaga 2280
Qy 2281 GGAATTAATGCACTTCTTATGATACCTGCTGACAGGAGATGATCTATGATGATCACTTC 2340
Db 2281 ggaatataatgcacttcttattgatatctgtgcagaggaatgatactgatagtgactctc 2340
Qy 2341 TAAAGAACTACACCCCAAGTGAAGAGAGTGAAGTAAATATGATGATGACGACAGCTC 2400
Db 2341 taaagaactacagccccaagtggaagaggtggaacaataaagtaagtcagtgaccagctc 2400
Qy 2401 CACCAATTCCTAGTCAAGTCAAGAGAAAGACCTAAATTAAGACAGAGACAAGTGAAGA 2460
Db 2401 caccattctcagtgctcaagtcagtaagaagaagccctaataatlaagacagagcaagtgaa 2460
Qy 2461 GGGATCTGGAGACTTGGATATATCTAGATGCTATTTGGTGATCTGATAGTTCTGACTT 2520
Db 2461 gggatctggagacttggatatatctagatgtatcttctgtgactgtgactgtgactgtgact 2520
Qy 2521 TTACAAATTAATTCATATCCCAATGGTATGATCTGAGGAGCTTAACCAACAGGTGTTCA 2580
Db 2521 ttacaataatccatatctcccaatggttagtcatctggggactcaagcaagatgtgttca 2580
Qy 2581 AGGAACATAATTCCTGGGTTGAAAAGTTGCACAGTCTGTGCACTTAATTCGCTCCATA 2640
Db 2581 aggaactaatatctctgtggtttgaaaagttcacagctgtgcagttatctgtctccata 2640
Qy 2641 TAACCGAGAGTGTCTGTGATAGCCCTGTTCTGTGGTCAAGTCCCTCCATTAATAAAA 2700
Db 2641 taacgagagtggtctgtgatagccctgttctgtgtgcaagctctccagtaaaaaa 2700
Qy 2701 TATCAGTGTCTTCCCATGTTACCAAAAGCAACCATGTTGGTGGGGAATCCAAAGATGAT 2760
Db 2701 tatcagtgcttctcccatgttaccaaagaacccaagttgtggtggaatccaaagatgtat 2760
Qy 2761 GGAATGACGAAATTAATGCTCAAGTATGAGTGGGCCAAACGAAATGATGACTGTGAC 2820
Db 2761 ggaatgacgaaatattatgtctcaagtaagtggttggtgccaacagaaatgtgactgtgac 2820
Qy 2821 TCAGACTCTTCTCTGAGAACTGAGGCTTACCAAACTCAAAAGCCGCGCAATGGAAC 2880
Db 2821 tcagactcttctctcaggaagctggtggttcaaaaactcaaaagccgagaaatgtgaacc 2880
Qy 2881 TATGAATTAATACTCATGGAAGACAGAGAGAGATTAATTAATCTTTTACCACGAC 2940
Db 2881 tatgaattcaaatccatcagtggaagacagagagatataataactcttaccacagacc 2940
Qy 2941 TGCATGTGGTGTCTATTCACATTTGCTCTTGTGATGATGATACACAGTGTGAG 3000
Db 2941 tgcattgtgggtgtctatttccacattgtccttctgtgtcctaataagataccaggtgcgag 3000

[illegible]

QY	4081	TCCTCCGAGAAATCCCATGATGCAACACCCGACGCTGCATCCATCTATACGTCTCAGA	4140
Db	4081	tcctcccgagaaatcccatgattgcaaacaccgaggtctgcatcattacatccagtcctccaga	4140
QY	4141	AATTGAAGGGCTGCGCATCTCGAAGAAATTTGGCCACGGAAACACTCTCTTTCCACAGCAGTT	4200
Db	4141	aattgaaggctgcgcatctcgaagaaatctggccaggaacagctccctctccagcagaagtt	4200
QY	4201	TCCCCACACAGGGGAAATCCTCGACAGTGTATGTATGTGGCGCATGAAATGCGACAGTGGTCA	4260
Db	4201	tggccacacaggggaaatcctcgacagtgtatgtatgtggtgcacatgaaatggaacagtggtca	4260
QY	4261	CATGGACAGATGAACATGAACACCCCATGTCTGCGCATGCCATAGGGTCTGTATCA	4320
Db	4261	catggacagatgaacatgaacaccccatgtgccatgtctgcatgacctatggtctccgtatca	4320
QY	4321	GAAATTACTGCTGACATCTCTGCAACGAGACCTTTTAAGGAACCACTGTATCAATATGACAC	4380
Db	4321	gaaatactgctgacatctctgcaacgagacctcttaaggaaacacctgtatacaatgacac	4380
QY	4381	TGCATGAGATTATTGGGAAGGAATCATTTGTCCAGACATCCATCTTGGAAAGAAAGACC	4440
Db	4381	tgcattgagattattgggaaggaaatcatctgtccaggcatccactctgtaagaaagaacc	4440
QY	4441	ACGTTTGACCTCCATCAAGGGTATTTTAAGTATGTCTATTGACGAGAATTTCTAG	4496
Db	4441	acgtttgacctccatcaagggtatcttaagtatgtctatttgacgagaaatcttag	4496
RESULT 3			
US-09-125-635-1			
: Sequence 1, Application US/09125635			
: GENERAL INFORMATION:			
: APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE			
: TITLE OF INVENTION: AIB1, A novel steriod receptor co-activator			
: FILE REFERENCE: 45944			
: CURRENT APPLICATION NUMBER: US/09/125,635			
: CURRENT FILING DATE: 1998-08-21			
: PRIOR APPLICATION NUMBER: 60/049,728			
: PRIOR FILING DATE: 1997-06-17			
: NUMBER OF SEQ ID NOS: 12			
: SOFTWARE: PatentIn Ver. 2.0			
: SEQ ID NO 1			
: LENGTH: 6835			
: TYPE: DNA			
: ORGANISM: Homo sapiens			
: FEATURE:			
: NAME/KEY: CDS			
: LOCATION: (201)..(4463)			
US-09-125-635-1			
Query Match 99.0%; Score 4451.6; DB 15; Length 6835;			
Best Local Similarity 99.5%; Pred. No. 0;			
Matches 4482; Conservative 0; Mismatches 9; Indels 15; Gaps 1.			
QY	1	GCTGATGATGACATCAGAGACCAATTAATAAATGCTTGACATCTCTTGACTGGTT	60
Db	116	gctgatactggtgaccccaagaccataataaactgcttgaaaccccttgactggtt	175
QY	61	AGCCAGTTCGTGATGTATATTCAAGTAGAGTGATGATTAGAGAAACTTGGATCCACTGGC	120
Db	176	agccagttgctgattatataccaagatggtgattagagaaacttgatccacttgc	235
QY	121	CAGTGATTCAGAAAACGAAATTGCCATGTGATACTCCAGACAAAGTCTTACTCGCAG	180
Db	236	cagtgattcacgaaaacgcaaaatgcacatgtatactccaagacaagaagctctaacctgcag	295
QY	181	TGGTGAAGAAAGGAGACGGGAGACAGAAAGTAAATATATTGAAGAAATTGGCTGACGTGAT	240
Db	296	tggtgaaaaacggaacggaagcaggaagaagtaataatatacttgaagaatctggctgaagctgac	355
QY	241	ATCTCCCAATCTTACTGATATTTGACAATTTCAATGTCAAAACCAAGATTAATGTGCGATTTT	300

|||||
Db 356 atctgcacatcttagttagtattgcacatttcaatgtcaacacagataatgtagcttt 415
OY 301 AAGGAACACGTAAGACAGATACGTCGAATTAAGAGACAGAGAAAAACTATTTCGAATGA 360
Db 416 aaaggaacacagataagacagatacgttcaataaagcaaggaataaactatttccatga 475
OY 361 TGATGATGTTCAAAAAGCGATGTATCTCTACAGGCGAGGAGTATGATTAAGACATC 420
Db 476 tgaatgtctcaaaaaagccgatgtattcttcaagggcagggagtattatgataaagatcc 535
OY 421 CTTAGACCGCTTTTACTTCAGGCAATGGATGTTTCTTATTTGTGGTGAATCGAGAGC 480
Db 536 cttagacgcgcttacttcaagcattgtagtattccatttggtagtgaatcagagag 595
OY 481 AAACATTGATTTGTATCAGAAAATGTCAACAATACCTGCAGATTAAGCAAGAGAGCCT 540
Db 556 aaacattgtatttgcatacgaataatgtcacacatacctgcataataagcaagagacct 655
OY 541 GGTAAACACAGTGTAAATATCTTACATGAGAGAGACAGAAAGGATTTCTTAAGAA 600
Db 656 ggttaacacagtgcttaacaatacttacaatgaagaagcagaagaagatttcttaagaa 715
OY 601 TTTACCAAAATCTACAGTTAATGAGATTCTCGACAAATGAGCCCAAGACAAAAAG 660
Db 716 ttaccaaactctacagttaattgaggttccctgcgaataatgacccaagaacaaaag 775
OY 661 CCATACATTTAATTTGCCGATGTGATGAAAAACACACATGATATTCTGGAAGACATAA 720
Db 776 ccatcatcttaattgcctgattgtatgtaaaacacacatactgatacttgaaagacataaa 835
OY 721 GCGCATCTGTAATGCGCCAGAGATATGAACAATGCAATGCACTTGCCCTGCTCAGCC 780
Db 836 cgcgcagctcgtgaatgycgcagagataatgaacaaatgcagtgcttgcctccatcagcc 895
OY 781 ACGACCTATGATGAGGAGGAGGAGATTGCAATCTTGATGATCTGTGTGACACGCC 840
Db 896 aagagctatgataagaggaaggaagattgtcaatctgtatgactgtgtagcagcgccg 955
OY 841 CATTCTACAGAGAGAAAGAACATTTCCATCAACACCTTGAGACTTTATTACAGACATGA 900
Db 956 catctcctaagagaaagacatttccataaaccttgagagcttattaccagacatga 1015
OY 901 TCTTCAGAGAAAGTTGTGCAATATAGATACAAATTCACGAGATCCTCCATGAGGCTCG 960
Db 1016 tcttccagagaaagtggtcaataatagatacaaatccactgagatctccatgaagccctg 1075
OY 961 CTTTGAAGATATATCCGAAGGTGATTCAGAGATTTTCTAGCTTAATGATGAGGACATC 1020
Db 1076 ctctgaagataatccgaaggtgattatcagagatttcttagtcttaaatgataatgtagcagtc 1135
OY 1021 ATGGTCCCAAGAAAGCTCACTATCAAGAGCTTATCTTAATGGCCATGCAGAAACCCAGT 1080
Db 1136 atggtccccaagaaacgtctcactatacaagaagcttataatgtagccatgcagaaccccaagt 1195
OY 1081 ATATGATTTCTGTTGGCTGATGGAATATAGTGAATGACTGCAGACAGAAAAAGAACTCTT 1140
Db 1196 atatcgatctcgtctgctgagatgagacatacagtgacgcagacaaaagcaaaactctt 1255
OY 1141 CCGAAATCCTGTACAAATGATGCATGGCTTTGTCTCAACCCATTCTCTCAGAGAGA 1200
Db 1256 ccgaatactctgtaacaataatgatacgacatgcttgcttcaaccactctcttcagagaga 1315
OY 1201 ACAGAAATGATATAGACCAAAACCCAAATCTCTTGACAGAGGATTAAGACACATATGCC 1260
Db 1316 acaagaatgagataaagacaaacccaatctctgtgacaagaagattagaaccaactatgac 1375
OY 1261 TGGATGCAACAGTTGGTGGGCGGATGATGTGCGCAAAACCAAGGTTTACAGATGCC 1320
Db 1376 tggatgcaacagttcgttagcgcgcagatgagatgtctgcgaacacagctcacaatgac 1435
OY 1321 GAGCAGACAGGCGCTATGGCTTGGCAGACCCCTAGCACACAGGCGAGATGATGAGCTAG 1380
|||||

Db 1436 gagcagcaagggcctatgcttgagcagaccctagcaaccacagggcagatgtagagctag 1495
OY 1381 GTATGGGGTTCACGTAACATAGCTTTCATTTGACCCCTGGGCGACGATTCGATCCACATC 1440
Db 1496 gtagtggggttccagtaataatagcttcatctgaccccttggccaaggaatgcaatcaatc 1555
OY 1441 TTCCTACACGAACAACACTATGGCTCAACTGATGATACCCCCCAATGAGAGTCTTGG 1500
Db 1556 ttccctacccaagaacaaataatgaggtcacaatgtagtggcccccaatggaggtccctg 1615
OY 1501 TCTTGCCCCAAACACAGCAGATATATCATGATTTCTCTGTAATCGTGGAGTCCAAAGAT 1560
Db 1616 tcttgcccaaaaacagcaagataatcatgatttctcctctgtaactgtaggagccaaagt 1675
OY 1561 AGCCTCACATGATTTTCTCTGTTGGAGGTGTGACACTCTCCATGGCATCTTCTGGCAA 1620
Db 1676 agcctcacatcagtttctcccttgtagagtgtagaacttcccatgtagctcttcgcaaa 1735
OY 1621 TACTGGGAACACAGCTTTTCCAGCAGCTCTCTCAGTGGCCCTGCAAGCCATCAGTGAAG 1680
Db 1736 tactgggaacacagcttctcagcagctctctcagtgccctgcgaagcatgaatgtagag 1795
OY 1681 TGTGGGACTTCCCTTTATCTACTCTGTATACACAGGCCCCAAATTTGATTAAGTCTTCC 1740
Db 1796 tgtgggacttcccttcttactctctgtcatctcacagggcccaaatgtagaactctcc 1855
OY 1741 CAATATGAATATTACCAACCAAGTAAGTAAGCATCAGATTCAGAGATTCCTGGG 1800
Db 1856 caatgaaataatcacaacccaagtaagtaagcaatcagattcacaagatcctctg 1915
OY 1801 CTTTTATTGCGACCAAAATCCAGTGAGAGTTCATATGTCTCAGTCAAAATAGCAGATGA 1860
Db 1916 ctttatttgcgacaaaataccagtgtagaggttcaatgttcaatgaataagcagatca 1975
OY 1861 CCTCAGTGCAAAAGAAAGTAAGAGAGCAGTGTGAGGGGCGAGAAATCAAGGGGTCC 1920
Db 1976 cctcagtgacaagaagaatgaagagagcagtgtagggggcagagaatacaaaagggctcc 2035
OY 1921 TTTGGAACCAAGGTCATAAAAAATTACTGAGTTACTTACTGTTCTTGATGATCCG 1980
Db 2036 ttggaagaacaaagtcataaaaatctactgagtaacttactctgtctcgtatgacg 2095
OY 1981 GGGTATTCCTCTTGACCAACCTCCCCCTATATTCAAATTCATTAAGAAATCTTGTTAG 2040
Db 2096 gggtaattctctcttgacccaactcccccttagatccaagtgtlaagaatcttctgtag 2155
OY 2041 TGTCAACAGCCCTCTGAGTCTCTCTCTATCTGTGAGGAGATATCTTATATCCAA 2100
Db 2156 tgtcaacagccctctgtagctctctctctacatctgtaggaatgatactctacatccaa 2215
OY 2101 TATGATGGGTCACTGTTCACAGAGAGACCGGATTTTGCAAGTGTCTGCAGAAATGG 2160
Db 2216 tatgataaggtactgttacaagagaagcacggaatttgcacaaggttgcagaaatg 2275
OY 2161 GAATTCACAGCTGAGGTGAGGCCAAGATTACTGCACAGGCACCTGGGAAAGACACAGCG 2220
Db 2276 gaattcacagctgaggtatgccaagattactgcagaagccacttggaaagaacacaaagc 2335
OY 2221 TATTAATCTTGTGGGAGCGAAATGTGTCAAGCAGAGCAGCTTAAGTCTTAAGAAGA 2280
Db 2336 tataactcttgtagggagcgaatgtgttcaagcagggcagctaaagctctaagaaga 2395
OY 2281 GGAAGATATGCACTTCTTAAGATACCTGCTGCAGACGGATGATCTAGTATGACATCTC 2340
Db 2396 ggaagataatgactctcttagataactctgtgacagggatgtagtcttagtgaatgac 2455
OY 2341 TAAAGAACTACAGCCCAAGGTGAAGGATGAGCAATAAATGAGACAGTGCACAGCTC 2400
Db 2456 taagaactacagccccaagtgtaggaatgataataatgagtcagtgacaagctc 2515
OY 2401 CACATTCCTAGCTCAAGTCAAGAGAGAACCCCTAAATTAAGACAGACAGCAAGTGAAGA 2460
Db 2516 cacatctcttagctcaagtcagaagaagaacctaataatlaagacagagacaagtgaga 2575
|||||

QY	2461	GGATCTCGAGACCTTGGATTAATCTACATGCTTTTCCTGGTACTGATCTATTCCTGACTT	2520
Db	2576	gggagcttcggagactcggataaaatccagaatgcatctccctcggatccgactagcttcgcgactt	2635
QY	2521	TTTCAATAATATTCATATATTCCTCAAAATGTGATGATCTGGGGATTAAGCAAGAGTGTTC	2580
Db	2636	ttacaataatctcataatctcccaaatggtatgcatctcggggactaaagaacagtgcttcca	2695
QY	2581	AGCAACTAATTTCTCTGGGTTTGAAAAGTTCCACAGTCTGTGCAAGTATTATTCCTCCATA	2640
Db	2696	aggaactcaatctctcgtggtcttgaaaagatccacagctctgtgcagctatctgctccata	2755
QY	2641	TAAACGAGCAGTGTGCTGGATATGACCCCTGTTTCGTGTGGCTCAACTCCTCCAGTAAAAA	2700
Db	2756	taaccggagcagtgctctctgataagccctgtctctgtgtctaaagtcctccaagtaaaaaa	2815
QY	2701	TATCAGTGTCTTCCCATGTTTACCAAAAGCAACCATGTTGGTGGGTAATGCCAAGATGAT	2760
Db	2816	tatcagtgctctcccccactgttaccaaaagaccccagtgtggttggaatccaaagaatgat	2875
QY	2761	GGATATGTCAGAGAAATTTATNGCTTCAGTATGTGGTGGGCCAAACGCAATGTGACTGTGAC	2820
Db	2876	ggaatagtcagaaataataatgctcaagataatggtggtgccaacacgaatgtaactgtgac	2935
QY	2821	TCAGACTCTCTTCCCTCAGAGAGACTGGGGCTTTACCAACTCAAAAGCCGGCACAATGTGAAC	2880
Db	2936	tcgactcctctctccaggaagactggggcttaccaaactcaagagccggcagaaatggaac	2995
QY	2881	TATCAATTTCAATCCATCCATGGGAAGACCAGAGAGAGATTATTAATCTTCTTTACCAAGAC	2940
Db	2996	tatgaatctcaaatcccaatggaagaccagagagagataataatctcttaccacagacc	3055
QY	2941	TGCACCTGGGTGGCTATTCCTCCACATTTGGCTCTTGCGTGTATATAGCATACAGAGTGACG	3000
Db	3056	tgcactgtgggtggctctatactcccaatctgcctcttcgtcttaataagcataaccaggtgtgag	3115
QY	3001	ACCAGTATTGCCAACACGACGACGACAGATGCTTCAATAGAGCCTGTGAATTCOCATGGG	3060
Db	3116	accagatattgcaaacagacagacagatgcttcaaaatgagccctggatgaatctcccatggtg	3175
QY	3061	AATGGGGGCTTAATTCCTCAATGAGCCACACAGAGATTCACAAATGGGTTTCCTGGCCGGA	3120
Db	3176	aaatgggggctcaatctcccaatgagccagacagacatactaaccaactcgggtctccgtgcgga	3235
QY	3121	TGGCATTGTTGCATGTGAAACAGTTTCTGATGACACTCAAAATAGACCCCTCTTCTTAGAA	3180
Db	3236	tggcatgttgtlcatgtgaacagttctctcatgtaactcaaaatagtgctctctcttagaa	3295
QY	3181	TTTCCCTGGATGATCTTGTGTGGGCCACTTTCACACTGGAAAGGCCAGATGACGAAAGAC	3240
Db	3296	ttccctggatgatacttgttctgtggccactcttcaaccctgtaaggtccagatgtgcgaagagtc	3355
QY	3241	ATTATTGACAGACGTGCACACTCTTCTTCAGCAAAACAGATGTGCAGAGGCTGTGAAGAAAT	3300
Db	3356	ataatgtgacagcagctgtaacactctctcttcgaagcaaacagatgcaacaagctcgtgaagaat	3415
QY	3301	TGCACAGACTTTGGGCACTTCTGCACTGTGCAATCAGGACAGCATTTAGAGCCCAACA	3360
Db	3416	tgcacagactcttgggcactctctctaacttgtlcaatcaggtacacagcatltagaccacaaca	3475
QY	3361	GGATGCTTTTCCAAAGGCCAACAACACACAGTAATGATGATGATCGAAGGAGGTTTATATAG	3420
Db	3476	ggatgctcttccaagccaaagaacagcagtaataatgataatgatacagaagcaggaattataatgg	3535
QY	3421	ACAGACATACCCACAGCACAGGGGCTTCGAATGCAAGAGAGCTTTCAATCTTTCAGGNCATTC	3480
Db	3536	acagaataaccacagcaacaggggtccccaatgcaaggaagcttcatcttccaggaacaatc	3595
QY	3481	ACCATCTTTTAACTCTATATGATATAGATGAACACGACAGGCAATTTTCTCTTCCAAAG	3540
Db	3596	accatctcttcaactataatgtatgaatcaagatgtaacccagcaaggaattctctctcccaag	3655

QY	3541	AATGACCCACGAGGCCAACTCATTTGAGACCCGGAAACAAACCCCAAGCAACTTAAAT	3600
Db	3566	aatgacccacgagccaaactcatTTGAGACCCGGAAACAAACCCCAAGCAACTTAAAT	3715
QY	3601	GCAGCTTCAGACAGGCTGCAGGGGCCAGCATTTTGAATCAGACGCACAGCACTTGA	3660
Db	3716	gcagcttcCagagggctgcagggccagcagtttttgaatCagacgcagacggcacttga	3775
QY	3661	ATTGAATAATGGAAAAACCTCTAGCTGGTGGTGGTGGGATGAGAGCCCTATGATGACACC	3720
Db	3776	attgaaataatggaaaacccctcagctggctggctggctggctggaTgagagccaaTgaTgacgc	3835
QY	3721	CCAGCAGGGTTTTCTTAATGCTCAATGTGTGCCCCAACGCACAGACAGACTGTAACTGA	3780
Db	3836	ccagcagggtttcttctaattgtcctaattgtgcgcccaagcagcagagagctgtctaTga	3895
QY	3781	TTCATTCCGACACAGAGGGGTGGCTTGTGATGATGCGACGAG-----CAGCA	3825
Db	3896	tcaattccgaaaaagagggtggtctatgataTgaTgacgacgacgacgacgacgaTga	3955
QY	3826	ACACACGACAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAACAGCAACAGCAGCAGCA	3885
Db	3956	gacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacga	4015
QY	3886	AACCCAGGCTTTCAGCCCACTGCTTAATGATGATGTTGCCACATGAGTGGGCTTTT	3945
Db	4016	aaccagggcttcagccacctcctaTgagctgtctcccaagcaTgagTgggcttctt	4075
QY	3946	GGCAGAGCCCAATGCCACAGACTCTCCGCAACAGTTTCCATTTCAACCAAAATTATG	4005
Db	4076	ggcagagcccaaatgccaagctctccgcaaaagtttccataTcaaccaaatatTg	4135
QY	4006	AATGGGACACAAACACATATGACGCTTTGGTGCAGTGTACTGCTCCCATGCAATGAT	4065
Db	4136	aatgggacaacaacacagatccagctcttgTcgaTgTcttccccaTgcaTgacTgac	4195
QY	4066	GTCGTCAGAGATGGGTCCCTCCAGATCCCATGATGATGACCAACCCGACAGCTGCATCAT	4125
Db	4196	gtcgTcaagaatTggTccctccagagatcccatTgatTgcaaacccgcaagctgtccat	4255
QY	4126	CTATACATGCTTCAGAAATGAAAGGGCTGGCCATAGGAATTTGGCAGAGACAGTCTCTT	4185
Db	4236	ctatcagctcccaagaaatgaaTgaggtctTgTcgaTcaagaaatctTgccaagaaacagctctt	4315
QY	4186	TTCCAGCAGCAGTGTTCGCCACACAGGGGATCCTCGAGTGTAGTATGTTGCACATGAA	4245
Db	4316	ttccagcagcaggtttgcccacagggaaTcctctgagTgataTgTgTgcacatTga	4375
QY	4246	TGCGACGAGTGTACATGTGGGACAGATGAACATGAACCCCATGCCATGTCTGGCATGCC	4305
Db	4376	tggcagcagTgtgtacaTgTgacagatTgaacaTgaaccccaTgccaTgtctTgTgaTgcc	4435
QY	4306	TATGGGTCTGATCAGAAATACGTGTGACATCTCTGCACACAGGACTCTTAAAGAAACCA	4365
Db	4436	tatggTctcgtatcagaaataTactgtcTgaactctctgcaccaggaactcttaagaaacca	4495
QY	4366	CTGTACAAATGACACTCTCACTAGGATTTATTTGGAGAGGATCATTTGTTCCAGCATCCATC	4425
Db	4496	ctgtacaatTgaacctctgactagatTatTtTgTgaagTaaTcatTgtcccaTgcaTcaTc	4555
QY	4426	TTGGAAGAAAGACACACTTTGAGCTTCATCAAGGCTATTTTAACTGATGTCATTTAGCC	4485
Db	4556	tTggaagaaagacacactTtTgagctTcaTcaagGctatTtTAActgATgTcattTtTgac	4615
QY	4486	AGGAAT 4491	
Db	4616	aggact 4621	

RESULT 4
US-60-068-511-1
; Sequence 1, Application US/60068511
; GENERAL INFORMATION:

APPLICANT: Suen, Chen-Shian
APPLICANT: Fraill, Donald E.
APPLICANT: Lytle, Richard C.
TITLE OF INVENTION: Cloning and Expression of a Nuclear
NUMBER OF INVENTION: Receptor Coactivator Proteins and Uses Thereof
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Home Products Corporation
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: USA
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/068,511
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Walsh, Andrea C.
REGISTRATION NUMBER: 34,988
REFERENCE/DOCKET NUMBER: 97243-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 973-683-2169
TELEFAX: 973-683-4117
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4789 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 185..4750
US-60-068-511-1

Query Match 98.6%; Score 4434.2; DB 38; Length 4789;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 4483; Conservative 0; Mismatches 8; Indels 24; Gaps 2;

QY 1 GCTGATGCTGACCTAGAGACCAATTAATTAACCTGCTTGAACATCCTTGACTGTT 60
DB 100 GCTGATGCTGACCTAGAGACCAATTAATTAACCTGCTTGAACATCCTTGACTGTT 159
QY 61 AGCAATGCTGATATATATTCAGATGATGATGAGGAAACTTGGATTCACCTGCG 120
DB 160 AGCAATGCTGATATATATTCAGATGATGATGAGGAAACTTGGATTCACCTGCG 219
QY 121 CAGTGAATCAGCAAAAGCAAAATTCATGATCTCCAGCAAGCTTCTTCACTGCG 180
DB 220 CAGTGAATCAGCAAAAGCAAAATTCATGATCTCCAGCAAGCTTCTTCACTGCG 279
QY 181 TGGTGAAGAAAGGAGAGGAGAGCAAGTAATATATTTGAAGAAATGGCTGAGCTGAT 240
DB 280 TGGTGAAGAAAGGAGAGGAGAGCAAGTAATATATTTGAAGAAATGGCTGAGCTGAT 339
QY 241 ATCTGCAATCTTAGTATGATGATGACAAATTCATGTCACCAACGATTAATTCGATTTT 300
DB 340 ATCTGCAATCTTAGTATGATGATGACAAATTCATGTCACCAACGATTAATTCGATTTT 399
QY 301 AAGGAAGAGTAAAGACAGATAGTCAATTAAGAGCAAGAAAGCAATATTTCCATGCA 360
DB 400 AAGGAAGAGTAAAGACAGATAGTCAATTAAGAGCAAGAAAGCAATATTTCCATGCA 459

QY 361 TGATGATGTTCAAAAAGCCGATGATCTTACAGGCGAGGAGTATGATTAAGACTC 420
DB 460 TGATGATGTTCAAAAAGCCGATGATCTTACAGGCGAGGAGTATGATTAAGACTC 519
QY 421 CTAGAGCCGCTTACTTACAGGCAATGATGATGCTTCTTATTTGCTGATGAGAGC 480
DB 520 CTAGAGCCGCTTACTTACAGGCAATGATGATGCTTCTTATTTGCTGATGAGAGC 579
QY 481 AAGCAATGATTTGATGATGAGAAATGTCACAAATACCTGCAATATTAACAGAGACT 540
DB 580 AAGCAATGATTTGATGATGAGAAATGTCACAAATACCTGCAATATTAACAGAGACT 639
QY 541 GGTAAACACAGGTTTAAATATCTTACATGAGAGACAGAAAGATTTCTTAAAGA 600
DB 640 GGTAAACACAGGTTTAAATATCTTACATGAGAGAGAGAGAGATTTCTTAAAGA 699
QY 601 TTTACCAAAATCTACAGTTAATGAGATTTCTGAGCAAAATGAGCCCAAAAGCAAAAAG 660
DB 700 TTTACCAAAATCTACAGTTAATGAGATTTCTGAGCAAAATGAGCCCAAAAGCAAAAAG 759
QY 661 CCATACATTTAATGCGGATGATGATGAGAAACACCAATGATATTTGGAAGACATAA 720
DB 760 CCATACATTTAATGCGGATGATGATGAGAAACACCAATGATATTTGGAAGACATAA 819
QY 721 CGCAGCTCGTGAATGCGCAGAGATATGAACCAATGAGCTTGGCCCTGACAGC 780
DB 820 CGCAGCTCGTGAATGCGCAGAGATATGAACCAATGAGCTTGGCCCTGACAGC 879
QY 781 ACAGCTATGATGAGAGAGGAGGAGATTTGCAATCTTATGATGCTGAGGACGCC 840
DB 880 ACAGCTATGATGAGAGAGGAGGAGATTTGCAATCTTATGATGCTGAGGACGCC 939
QY 841 CATTAACAGAGAGAAAGCAATTTCCATCAAACTGAGAGCTTATTTACAGACATGA 900
DB 940 CATTAACAGAGAGAAAGCAATTTCCATCAAACTGAGAGCTTATTTACAGACATGA 999
QY 901 TCTTTAGAGAAAGGTGCAATATATGATCAAAATTCACAGATCTCCATAGAGGCTG 960
DB 1000 TCTTTAGAGAAAGGTGCAATATATGATCAAAATTCACAGATCTCCATAGAGGCTG 1059
QY 961 CTTTGAAGATATATATCGAAGGTGATTCAGAGATTTTATGCTTAATGATGAGGACATC 1020
DB 1060 CTTTGAAGATATATATCGAAGGTGATTCAGAGATTTTATGCTTAATGATGAGGACATC 1119
QY 1021 ATGCTCCAGAAAAGTCACTATTCAGAACTTATCTTAAATGGCCATGAGAAAACCCAGT 1080
DB 1120 ATGCTCCAGAAAAGTCACTATTCAGAACTTATCTTAAATGGCCATGAGAAAACCCAGT 1179
QY 1081 ATATGATTTCTGCTTGGCTGATGAGAACTATAGTACTGACAGACAAAAGCAAACTCTT 1140
DB 1180 ATATGATTTCTGCTTGGCTGATGAGAACTATAGTACTGACAGACAAAAGCAAACTCTT 1239
QY 1141 CCGAAATCCGTAAACAAATGATGACATGCTTGTCTCAACCCATCTTCTCAGAGAGA 1200
DB 1240 CCGAAATCCGTAAACAAATGATGACATGCTTGTCTCAACCCATCTTCTCAGAGAGA 1299
QY 1201 ACAGATGATATAGACCAAAACCAATCCCTGTTGGACAAAGGATTAACCACTATGCG 1260
DB 1300 ACAGATGATATAGACCAAAACCAATCCCTGTTGGACAAAGGATTAACCACTATGCG 1359
QY 1261 TGGATGCAACAGTTCGCTGAGGCGCATGAGATGTCGCCAACAAGCTTACAGATGCC 1320
DB 1360 TGGATGCAACAGTTCGCTGAGGCGCATGAGATGTCGCCAACAAGCTTACAGATGCC 1419
QY 1321 GAGCAGCAGGCGCTATGCTTGGCAGACCTGACACACAGGCAATAGAGGAGCTAG 1380
DB 1420 GAGCAGCAGGCGCTATGCTTGGCAGACCTGACACACAGGCAATAGAGGAGCTAG 1479
QY 1381 GTATGGGGGTTCCAGTAAATAGCTTCAATGACCCCTGGGCCAGGAGATCAATCAACATC 1440
DB 1480 GTATGGGGGTTCCAGTAAATAGCTTCAATGACCCCTGGGCCAGGAGATCAATCAACATC 1539

1441 TTCTACCAAGAACAACTATGGCTCAACATGATGAGTACGCCCCCAATGGAGTCTGG 1500
1540 TTCTACCAAGAACAACTATGGCTCAACATGATGAGTACGCCCCCAATGGAGTCTGG 1599
1501 TCTTCCCCCAACAGCAAGATATCATGATTTCTCTCTGTAATGCGGGAGTCCAAAGAT 1560
1600 TCTTCCCCCAACAGCAAGATATCATGATTTCTCTCTGTAATGCGGGAGTCCAAAGAT 1659
1561 AGCCTACATCATGTTTCTCTGTTGAGGTGACATCTCTCCATGGCATCTTCTGGCAA 1620
1660 AGCCTACATCATGTTTCTCTGTTGAGGTGACATCTCTCCATGGCATCTTCTGGCAA 1719
1621 TACTGGGAACCAAGCACTTTTCCAGAGCTCTCTGAGTGGCCCTGCAAGCATAGTGAAG 1680
1720 TACTGGGAACCAAGCACTTTTCCAGAGCTCTCTGAGTGGCCCTGCAAGCATAGTGAAG 1779
1681 TGTGGGAGCTTCCCTTTTATCTACTCTGTCATCACCAGGCCCCCAATTGGATACCTCC 1740
1780 TGTGGGAGCTTCCCTTTTATCTACTCTGTCATCACCAGGCCCCCAATTGGATACCTCC 1839
1741 CAATATGAATATTTACCAACCAAGTAAAGTAAGCAATCAGATTTCCAAAGTCTCTGGG 1800
1840 CAATATGAATATTTACCAACCAAGTAAAGTAAGCAATCAGATTTCCAAAGTCTCTGGG 1899
1801 CTTTATTTGCCACCAAAATCCAGTGGAGAGTTCAATGTGTCTGATCAATATAGCAGATCA 1860
1900 CTTTATTTGCCACCAAAATCCAGTGGAGAGTTCAATGTGTCTGATCAATATAGCAGATCA 1959
1861 CCTCAGTACCAAGAAATAGAGAGAGTGTGAGGGGCGAGAAATCAAAAGGGTCC 1920
1960 CCTCAGTACCAAGAAATAGAGAGAGTGTGAGGGGCGAGAAATCAAAAGGGTCC 2019
1921 TTTGGAAGCAAAAGTCAATAAAATTTACTGAGTTACTTACTCTTCTCTGATGACCG 1980
2020 TTTGGAAGCAAAAGTCAATAAAATTTACTGAGTTACTTACTCTTCTCTGATGACCG 2079
1981 GGGTATTTCTCTTGTACCAACTCCCTCTAGATTCAAGTTGTAAGAAATCTTCTGTAG 2040
2080 GGGTATTTCTCTTGTACCAACTCCCTCTAGATTCAAGTTGTAAGAAATCTTCTGTAG 2139
2041 TGTCAACGAGCCCTCTGAGAGTCTCTCTCTACATCTGGAGAGATATCTCTACATCCAA 2100
2140 TGTCAACGAGCCCTCTGAGAGTCTCTCTCTACATCTGGAGAGATATCTCTACATCCAA 2199
2101 TATGATGGGTCTCTGTTTACAAAGAGACAGCCGGATTTTGCCAAGTTGTCTGAGAAATGG 2160
2200 TATGATGGGTCTCTGTTTACAAAGAGACAGCCGGATTTTGCCAAGTTGTCTGAGAAATGG 2259
2161 GAATTCACACAGCTGAGTATGAGCAAGATTACTGCAACAGGCACTGGGAAAGACACCCAGAG 2220
2260 GAATTCACACAGCTGAGTATGAGCAAGATTACTGCAACAGGCACTGGGAAAGACACCCAGAG 2319
2221 TATAACTCTTGTGGGGAGCGGAATGTTGTCAAGCAGAGCAGCTTAACTCTTAAAGAA 2280
2320 TATAACTCTTGTGGGGAGCGGAATGTTGTCAAGCAGAGCAGCTTAACTCTTAAAGAA 2379
2281 GGAGAAATATGACATCTTCTTAATACCTGCTGACAGGGATGATCTTACTGATGACCTCTC 2340
2380 GGAGAAATATGACATCTTCTTAATACCTGCTGACAGGGATGATCTTACTGATGACCTCTC 2439
2341 TAAAGAACTACAGCCCAAGTGAAGAGATGGAGCAATAAATGATCTAGTGACACAGCTC 2400
2440 TAAAGAACTACAGCCCAAGTGAAGAGATGGAGCAATAAATGATCTAGTGACACAGCTC 2499
2401 CACCAATTCCTAGCTCAAGTCAAGAGAAAGACCTTAAATTAAGACAGAGCAAGTGAAGA 2460
2500 CACCAATTCCTAGCTCAAGTCAAGAGAAAGACCTTAAATTAAGACAGAGCAAGTGAAGA 2559
2461 GGGATCTGAGACTTGGATTAATCTAGATGTAATCTTGGTATCTGACTAGTCTGACTT 2520
2560 GGGATCTGAGACTTGGATTAATCTAGATGTAATCTTGGTATCTGACTAGTCTGACTT 2619
2521 TTACAATAATTCATATCTCTCAATGATGATCATCTGGGGACTAAGCAACAGCTTTTCA 2580

2620 TTACAATAATTCATATCTCTCAATGATGATCATCTGGGGACTAAGCAACAGCTTTTCA 2679
2581 AGCACTAATTCCTGCGGTTTGAAGAGTTTACAGCTGCTGGAGTCTATTCGCTCCATA 2640
2680 AGCACTAATTCCTGCGGTTTGAAGAGTTTACAGCTGCTGGAGTCTATTCGCTCCATA 2739
2641 TAACGAGCAGTGTCTCTGATAGGCCGTCTTCTGTTGGCTCAAGTCTCCAGTAAAAA 2700
2740 TAACGAGCAGTGTCTCTGATAGGCCGTCTTCTGTTGGCTCAAGTCTCCAGTAAAAA 2799
2701 TATCAGTGTCTTCCCATGTTTACCAAGCAACCCATGTTGGTGGGAATCCAAAGATAT 2760
2800 TATCAGTGTCTTCCCATGTTTACCAAGCAACCCATGTTGGTGGGAATCCAAAGATAT 2859
2761 GATATGTCAGGAAATTTATGGCTCAAGTATGGGTGGGCAACCGAAATGTGACTGTAC 2820
2860 GATATGTCAGGAAATTTATGGCTCAAGTATGGGTGGGCAACCGAAATGTGACTGTAC 2919
2821 TCAGACTCTCTCTCAGAGAGACTGGGGCTTACCAAACTCAAAAGGCCGAGAAATGGAAC 2880
2920 TCAGACTCTCTCTCAGAGAGACTGGGGCTTACCAAACTCAAAAGGCCGAGAAATGGAAC 2979
2881 TATGAATTCAAACTCCATGGGAAAGACAGAGAGATTTATATCTTTTACCAGACC 2940
2980 TATGAATTCAAACTCCATGGGAAAGACAGAGAGATTTATATCTTTTACCAGACC 3039
2941 TGCACCTGGGTGCTTATTTCCCATGCTCTCTGCTTATATAGCATACAGAGTGGCAG 3000
3040 TGCACCTGGGTGCTTATTTCCCATGCTCTCTGCTTATATAGCATACAGAGTGGCAG 3099
3001 ACCAGTATTTGCAACAGCAGCAGAGATCTTCAAAATGAGGCCGTGTGAATCCCATGGG 3060
3100 ACCAGTATTTGCAACAGCAGCAGAGATCTTCAAAATGAGGCCGTGTGAATCCCATGGG 3159
3061 AATGGGGCTATATCCCTTATGGGCAAGCAGAGCATTAACAACATGGGCTTCTGGCCCCGA 3120
3160 AATGGGGCTATATCCCTTATGGGCAAGCAGAGCATTAACAACATGGGCTTCTGGCCCCGA 3219
3121 TGGCATGTTGTCCATGGAACAAATTTCTCATGAGCACTCAAAATAGGCCCTTCTTAGGAA 3180
3220 TGGCATGTTGTCCATGGAACAAATTTCTCATGAGCACTCAAAATAGGCCCTTCTTAGGAA 3279
3181 TTTCCCTGATGATCTTGTGTTGGCCACCTTCCAACTGGAAGGCCAGAGTGAAGAGC 3240
3280 TTTCCCTGATGATCTTGTGTTGGCCACCTTCCAACTGGAAGGCCAGAGTGAAGAGC 3339
3241 ATTATTTGACAGCTGACACACTCTTCTCAGCAACAGATGCGACAGGCTTGAAGAAAT 3300
3340 ATTATTTGACAGCTGACACACTCTTCTCAGCAACAGATGCGACAGGCTTGAAGAAAT 3399
3301 TGCAGAGCTTTGGGCACTTCTGAACTTGTCAATCAGAGGAGAGCATTAAGAGCCAAACA 3360
3400 TGCAGAGCTTTGGGCACTTCTGAACTTGTCAATCAGAGGAGAGCATTAAGAGCCAAACA 3459
3361 GSATGCTTTTCCAAAGGCCAAAGAGCAGAGTATGATGATGAGAGGAGCATTAATATGG 3420
3460 GSATGCTTTTCCAAAGGCCAAAGAGCAGAGTATGATGATGAGAGGAGCATTAATATGG 3519
3421 ACAGCATACCCAGCAGAGGGGCTCCAAATGCAAGGAGAGCTTATCTTCAGGGCAATC 3480
3520 ACAGCATACCCAGCAGAGGGGCTCCAAATGCAAGGAGAGCTTATCTTCAGGGCAATC 3579
3481 ACCATCTTTTAACTCTATGATGAATCAGATTAACAGCAAGCAATTTTCTCTCCAAAG 3540
3580 ACCATCTTTTAACTCTATGATGAATCAGATTAACAGCAAGCAATTTTCTCTCCAAAG 3639
3541 AATGCAACCCAGCAGCAACATCATATGAGACCCCGGCAAAACACCCCAAGCAACTTAGAAT 3600
3640 AATGCAACCCAGCAGCAACATCATATGAGACCCCGGCAAAACACCCCAAGCAACTTAGAAT 3699
3601 GCAGCTTACGAGAGAGGCTGACAGGGGCAAGCTTTTGAATACAGCCGACAGGCACTTTGA 3660

Db	3700	GCACCTTCACAGAGGGCTGCAGGGCCAGAGATTTTGAATCAGAGCCGCAAGCACTTGA	3755
Oy	3661	ATTGAAATGGAAAAACCTTACTCTGTGTGTGCTGCGGTGATGAGGCTTATGATGACCC	3720
Db	3760	ATTGAAATGGAAAAACCTTACTCTGTGTGTGCTGCGGTGATGAGGCTTATGATGACCC	3819
Oy	3721	CCAG-----CAGGCTTCTTAAATGCTCAAAATGGTGGCCAAAGCAGAGAGACT	3771
Db	3820	CCAGCTGAGCTCCGAGGGTTTCTTAATGCTCAAAATGGTGGCCAAAGCAGAGAGACT	3879
Oy	3772	GCTAAGTCATCACTTCCGACACAGAGGGGTGCTATGATGATGACGACGACGACAA-----	3826
Db	3880	GCTAAGTCATCACTTCCGACACAGAGGGGTGCTATGATGATGACGACGACGACGACAA	3939
Oy	3827	-----CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACGACAAAGCAACAGCAGCA	3876
Db	3940	GCAACAGCAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAACGACAAAGCAACAGCAGCA	3999
Oy	3877	ACAGACGAAACCCAGGCTTTCAGCCCACTCTTAATGTACTGCTTCCCCAGCATGGA	3936
Db	4000	ACAGACGAAACCCAGGCTTTCAGCCCACTCTTAATGTACTGCTTCCCCAGCATGGA	4059
Oy	3937	TGGGCTTTTGGCAGAGACCCCAATGGCACAAGTCTCTCGGCAACAGTTTCCATATCAAC	3996
Db	4060	TGGGCTTTTGGCAGAGACCCCAATGGCACAAGTCTCTCGGCAACAGTTTCCATATCAAC	4119
Oy	3997	AAATATGAAATGGAGACAACAACCAATCAGCTTGGTTCAGTGTAGTCTCCCAA	4056
Db	4120	AAATATGAAATGGAGACAACAACCAATCAGCTTGGTTCAGTGTAGTCTCCCAA	4179
Oy	4057	TGCATGATGTCTGCAAGAAATGGGTCTCTCCCAATTCATGATGACAAACCCGACAGC	4116
Db	4180	TGCATGATGTCTGCAAGAAATGGGTCTCTCCCAATTCATGATGACAAACCCGACAGC	4239
Oy	4117	TGCATTCATCTATAGTCCCAAGAAATGAAAGGCTGGCCATGAGAAATTTGGCCAGGAA	4176
Db	4240	TGCATTCATCTATAGTCCCAAGAAATGAAAGGCTGGCCATGAGAAATTTGGCCAGGAA	4299
Oy	4177	CAGCTCCTTTTCCGAGCAGCAGATTTGCCCAACAGGGAATCTCTGCAAGTATATGATGAT	4236
Db	4300	CAGCTCCTTTTCCGAGCAGCAGATTTGGCCCAACAGGGAATCTCTGCAAGTATATGATGAT	4359
Oy	4237	GCACATGAATGGCAGCAGTGGTCAATGGGACAGATGAACATGAACCCCATGCCATGTC	4296
Db	4360	GCACATGAATGGCAGCAGTGGTCAATGGGACAGATGAACATGAACCCCATGCCATGTC	4419
Oy	4287	TGGCATGCTATGGGTGCTCTATCAGAAATATCTCTACATCTCTGACACGAGACTCTTA	4356
Db	4420	TGGCATGCTATGGGTGCTCTATCAGAAATATCTCTACATCTCTGACACGAGACTCTTA	4479
Oy	4357	AGGAAACACACTGTCAAAATGACACTGACATGAGATTTATTTGGAGAAGAAATCATTTGTTCCAG	4418
Db	4480	AGGAAACACACTGTCAAAATGACACTGACATGAGATTTATTTGGAGAAGAAATCATTTGTTCCAG	4539
Oy	4417	GCATCCATCTTGAAGAAAGGACAGCTTGAAGCTCCATCAAGGATATTTAAGTATGT	4476
Db	4540	GCATCCATCTTGAAGAAAGGACAGCTTGAAGCTCCATCAAGGATATTTAAGTATGT	4599
Oy	4477	CATTGAGCAGGAAT 4491	
Db	4600	CATTGAGCAGGAAT 4614	
RESULT 5			
US-09-359-922-3962			
Sequence 3962, Application US/09359922			
GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc.			
; TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA			
; FILE REFERENCE: 20411-752CON1			
; CURRENT APPLICATION NUMBER: US/09/359,922			
; CURRENT FILING DATE: 1999-07-22			

```

? EARLIER APPLICATION NUMBER: US 09/205,155
? EARLIER FILING DATE: 1998-12-03
? NUMBER OF SEQ ID NOS: 13203
? SOFTWARE: FastSeq for Windows Version 3.0.0
? SEQ ID NO 3962
? LENGTH: 6855
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(6855)
? OTHER INFORMATION: n = A,T,C or G
US-09-359-922-3962

```

Query Match	98.58;	Score 4429.6;	DB 17;	Length 6855;
Best Local Similarity	99.28;	Pred. No. 0;		
Matches 4482; Conservative	0;	Mismatches	9;	Indels 27; Gaps 2

QY	1	GGGGATGTGGTACGACACAGACGCAATAAAATTAACGCTTGAAACATCCCTTGACGTG	60
DB	116	gcttggatctggctgaccacagagaccataaaataaaccttgcttgaacatcccttgactcgtt	175
QY	61	AGCCAGTTGGCGATGTATATTTCAGATGAGAGGATTAAGGAACCAACTTGGATCCACGCG	120
DB	176	agccagttcgtcgtgattatattccaagatgagctgtggttttgagaaaacttgcattccactcgc	235
QY	121	CAGTGATTTACGAAAAACGCAAAATTTGCCATGTGATACTCAGAGCAAGGTCTTACCTCAG	180
DB	236	caagtattccagaaacgcaatttgcattgcatctgtaaccacagacaagctcttaccctgcag	295
QY	181	TGCTAAAAAAGGAGACGCGGAGCAGAGAAAGTATATATTGAGAAATTTGGCTGACGTGAT	240
DB	296	tggctaaaaacggaacggaacggaagaaagtaataatactatgaactcttgctgcgtcgtat	355
QY	241	ATCTCCCATCTCTTACTGATTTTGGCAATTTGCAATTCAGTCAAAACACATTAATGTGGATTTT	300
DB	356	atctcccatctcttactgattttgcaatttgcaatttcaatgtaaccagataaactgtgcgatttc	415
QY	301	AAAGGAACAGTAAACAGACGATACGTCAATTAATAAGACAAAGAAAACATATTTCCAAATGA	360
DB	416	aaaggaacacgtaagacacgatactgctcaataaagagcaagaaaactatcttccaatga	475
QY	361	TGATGATGTTCAAAAAGCCGATGATATCTTCTTACAGGCGAGGAGCTATTGATTAAGACTC	420
DB	476	tgatgattgttcaaaaagccgattgtctcttccacaggaaggttatgtataaagatc	535
QY	421	CTTAGGACCCGCTTTACTTTCAGGCAATTGGATGTTTCTATTGTGGTGAATGAGAGGC	480
DB	536	cttaggaccgccttactctcaggaattggaatggttctctatttggltgtaatcgaacg	595
QY	481	AAACATTTGATTTGATCAGAAAATGTCCACCAATTACCTGCATATTAAGCAAGAGACCT	540
DB	596	aaacattgtatttgcatacagaaaaatgtccacaataactctgcaatatagaacgaagacct	655
QY	541	GGTTAAACACAGCTGTTTACAATATCTTACATGAAGAAACAGAAAGGATTTTCTTAAGAA	600
DB	656	ggttaacacagtgcttaacaatattctactatgatgaagacgaagaagatttcttaagaa	715
QY	601	TTTACCAAAATCTACAGTTAATGAGATTCTCTGGCAAAATAGCCCCAAGACAAAAAAG	660
DB	716	tttaccaaaattcaacagttaatggaatttctctgcgaacaatlgagaccacaagaacaaaag	775
QY	661	CCATTCATTTAATTCGCGATGTTGATGAAAAACACCATATATTTCTGGAAGACATAAA	720
DB	776	ccattcaatttaattcgtatgtaatgtaagaacaccacaatgatacttcggaagacataaa	835
QY	721	CGCCAGTCTGAAATGCGCCACAGAGATATGAACAATGCAAGTCTTTGCCCTGTCCAGCC	780
DB	836	cgcagctctctgaatatgcgcagagatgataaacaatlgagtgcttgcctgtctcagcc	895
QY	781	ACGAGCTATGATGAGGAAGGGGAAGATTTCGAATCTTGATGATCTGTGTGGCAGCCG	840

Db 896 acgagctatgatgagggaaagatttgcacattctgtatgatctgtgtgcaagccg 955
QY 841 CATTCTACAGAGAAAGACATTTTCATCAAACTCGAGCTTTATTACAGACATGA 900
Db 956 cacttactcagaagaaagacaattccatcaaaccttgagagcttattacacagatga 1015
QY 901 TCTTCAGAAAGGTGTTGATATATAGATACAAATTCAGTGAATCTCCATGAGGCTGG 960
Db 1016 tcttcccaaaaaggtctgtccaaatagatacaaatccagagatccccaagagccctgg 1075
QY 961 CTTTGAAGATATATATCCGAAGGTGATTCAGAGATTTTCTTAAATATGATGGCAGTC 1020
Db 1076 ctttgaagataataaccggaaggtgtatcagaagatttctttagtccaatagatgggcagtc 1135
QY 1021 ATGTGTCACGAAAGCTCACTATCAAGAGCTTATCTTAATGGCCATGCAAAACCCCACT 1080
Db 1136 atggtccccaagaaacgtctcactatcaagaagcttatcttaatggtccatgycgaagaccacagtc 1195
QY 1081 ATATGATTTCTGTTGGCTGATGGAACATATAGACTGACAGACAAAAGCAACTCTTT 1140
Db 1196 atatcgattctctgtgtcgtatggaacctatgtacatgcacagacaaaagcaacctctc 1255
QY 1141 CCGAAATCCTGTACAAATGATGCAATGGCTTTGTCTCAACCCACTTCTCTCAGAGAGA 1200
Db 1256 ccgaaatccctgtaacaaatgatalcgacatggtcttgttccaacccaactctctcagagaga 1315
QY 1201 ACAGAAATGATATACCAACCAAACTCTTGTGGACAGAGGATTTAGACCACTATGGC 1260
Db 1316 acagaaatgataataagcaaaaacccaatccctgttgtacaagaagatataagaccctatggtc 1375
QY 1261 TGGATGCAACAGATTGCGTAGCGCGCATGATGATGTGCGCAACCAAGGCTTACAGATGCC 1320
Db 1376 tggatgcaaacagttctcgtgtcaggtcaggtatgtatgtgcgccaaccaaagcttccagatgccc 1435
QY 1321 GAGCAGCAAGGCGCTATGGCTTGGCAGACCTTAGCACACACAGGCGAGATGAGTGAAGTCC 1380
Db 1436 gagcagcagggcctatggtccttgcagaccctagcaccaacagggcagatgagtgaggagctag 1495
QY 1381 GTATGGGGGTTTCCATACATATAGCTTCAATGACCCCTGGGCCAGGCAATGCAATCACATC 1440
Db 1496 gtaatgggggttccagtaacatagcttcatctgaccccttggccaaggtcagatccatccatcc 1555
QY 1441 TTCCTACAGAGAACACTATGGGCTCAACATGAGTAGGCCCCCAACATGGGAGTCCGG 1500
Db 1556 ttccctacagaacaaacaactatgggtctcaacatgagtagccccccaagagatccctgg 1615
QY 1501 TCTTGGCCCCAAACAGCAGAAATATCATGATTTCTCTCGTAAATCGTGGGAGTCCAAAGAT 1560
Db 1616 tcttggccccaaacagcagaataatcatgattctctcctcgttaatgtgtggagtcacaagat 1675
QY 1561 AGCCTACATCAGTTTCTCTCTGTGGCAGGTGTGCACCTCTCCATGGCATCTTCTGGCAA 1620
Db 1676 agccctacatcaagttcttctcctgtgcaggtgtgcactctcccatgycatcttctcggaa 1735
QY 1621 TACTGGGAACACAGCTTTTCCAGCAGCTCTCAGTGCCCTCGCAAGSCATCAGTGAAG 1680
Db 1736 tactgggaacacagcttctccagagctctctcaggtgcctcgcagacatcaagtgaaag 1795
QY 1681 TGTGGGGACTTCCCTTTATCTACTCTGTCAATCACAGGCCCAATTTGATTAATCTTCC 1740
Db 1796 tgtggggacttccctcttattactactcgtgtcatcaacaggtccccaattgtgataaactctcc 1855
QY 1741 CAATATGAATATATACCAACCAAGTAAGTAAGCAATCAGATTTCAGAGTCTCTGGG 1800
Db 1856 caatataatattaccacaacaaagtaagtaagcaatcaagatttccaagagtcctctgg 1915
QY 1801 GTTTATTTGGCGCAAAATCCAGTGGAGAGTTCAATGTGTCAATGTCAATATAGACAGATCA 1860
Db 1916 ctttataatgcaacaaatccaggtgagagttcaaatgtcagttcaaaataagcagaagatca 1975
QY 1861 CCTCAGTACAAAGAAATTAAGAGAGCAGTGTGAGGGGCGAGATCAAAAGGGGTCC 1920
Db 1976 cctcagtgacaaagaaagtaagagagcaggtgtgtgaggggcagagaaatcaaaaggtctcc 2035

QY 1921 TTTGCAAAAGCAAGGTCAATAAAAAATTACTGCAGTTACTTACCTGTTCTTGATGACCG 1980
Db 2036 ttctggaagcaaaaggtctataaaaaattactcagttacttactctgtctctcgtacgacg 2095
QY 1981 GGGTATTCCTCTTGTACCAACTCCGCCCTGATTCAGAGTTGTAAAGAAATCTTCTGTAG 2040
Db 2096 ggggtatctctctcttgacccaactcccccttgatattcaagttgttaagaatcttctgttag 2155
QY 2041 TGTCAACAGCCCTCTGAGAGTCTCCTCTCTCATCTGTGAGAGATATCTGTACATCCAA 2100
Db 2156 tgtcaacagccccctctgtagtctctctctctctctctctcagtgagagatatactctcaatccaa 2215
QY 2101 TATGCATGGGCTGACTGTTTACAAGAGACAGCAGATTTTGCACAAATGCTGACAGATGG 2160
Db 2216 tatgcatggtctactgattacaagagaagcacccgattcttgacaaagttcgtgcagaatcgg 2275
QY 2161 GAATTCACAGCTGAGGTAGCCCAATTAATACAGCAACGCCTGGGAAAGACACAGAG 2220
Db 2276 gaattcacagctgaggttagccaagattacgtcagaagccacgtgggaaagacaacagcag 2335
QY 2221 TATAACTTCTTGTGGGACGGAATGTTGTCAAGCAGAGAGCAGTAACTCTTAAGAAAG 2280
Db 2336 tataacttctgttgggagcggaaatgtgttcaagagaaggaagctaaagttcctaagaagaa 2395
QY 2281 GGAGAAATATGCACTTCTTACATACCTGCTGACAGGATGATCTTACTGATGCACTCTC 2340
Db 2396 ggaagataatgacttctcttagatctactctgtgacagggatgataccctagtgatgcactctc 2455
QY 2341 TAAAGACTACAGCCCCCAAGTGGAGGAGTGGACAAATTAATGATGCTGACACAGCTC 2400
Db 2456 taagaactacagcccccaagtggaaggaatggataataaagaaatcagtgaccaagcttc 2515
QY 2401 CACCATTCCTACCTCAAGTCAAGAGAAAGACCTTAATTAAGACAGAGCAAGTGAAGA 2460
Db 2516 caccatctctcgtcgaagttcaagagaagacccctaaatataagaaagaaagaaagaa 2575
QY 2461 GGGATCTGGAGACTTGGATTAATCTTAGATGCTAATTTTGGTGTATCTGACTTCTGACTT 2520
Db 2576 gggatctggagacttgagataacttagatgactatctctgtgtgactgactagttctgcactt 2635
QY 2521 TTACAAATTAATTCATATCTCAATGATGATGATGATGATGATGATGATGATGATGATG 2580
Db 2636 ttacaataatccataatccataatccataatccataatccataatccataatccataatccata 2695
QY 2581 AGGAACATAATCTCTGGGTTTGAAGAGTTTCAACAGTCTGTGAGTCTATCTCTCTCATTA 2640
Db 2696 aggaactaattctctcgtgttgaagagttcaacagttcgtgcagttctatctgtctccata 2755
QY 2641 TAACCGACAGTGTCTGTGATAGCCCTGTGTTGTGTTGGTCAAGTCTCTCAAGTAAAAA 2700
Db 2756 taaccgagcagtgctctcgtgataagccctgtctgtgtgcataagttctccagtaaaaaa 2815
QY 2701 TATCAGTGTCTTCCCGATGTTACCAAAACCAATGTTGGGTGGGAATCCAAAGATGAT 2760
Db 2816 taccagtgcttctcccaatgattacaagaacccaatglttgggtgggaatcccaagaatgat 2875
QY 2761 GGATAGTACGAGAAATTAATGCTCAAGTATGGGTGGGCGCAAAACGAAATGTGACTGTGAC 2820
Db 2876 ggaatagttcaagaaaaatttgtctcaagttatgtgtgggtccaaacccaagatgtgactgtgac 2935
QY 2821 TCAGACTCTCTTCTTCAGAGAGACTGGGGCTTACCAAACTCAAAAGCCCGGACAGATGGAAC 2880
Db 2936 tcagactctctctctcaagagacttgggttaccaaactcaaggtccggcagaatgagaaac 2995
QY 2881 TATGAATTCAAATTCATAGGGAAGACGAGAGAGATTAATTAATCTTCTTACCAAGACC 2940
Db 2996 tatgaattcaaatccatcagtggaagaccaggtgagatataacttcttaccagaaacc 3055
QY 2941 TGCACTGGGTGGCTTATTCACATTCGCTTGTGGGTCAATGCAATACAGGTGGAG 3000
Db 3056 tgcactgggtggtctctatctcccaacttgcctctcgtcttaataagcatacaaggtcgcag 3115

|||||
Db 236 cagtgatccagaaaaagcaaatgycalgtgatactccaggaacaggtcctacgtcag 295
QY 181 TGTGTAAAAAGGAGACGGAGCAGAGAAAGTAAATATATTGAAGATTGGCTGACCTGAT 240
Db 296 tggtagaaaaagcgagcgagcgagcaagaataatataatgaaagatctgycgtgagctgat 355
QY 241 ATCTGCCAATCTTACTGATATTGACAAATTTCAATGTCAACAGACAGTAAATGTGCAATTTT 300
Db 356 atctgcgaatcttagtgatattgacaattccaatgtccaacacagaataatgtgcatttt 415
QY 301 AAGGAACAGTAAGACAGATACGTAATTAAGAGAGAGAAAGAACTATTTCGAATGA 360
Db 416 aagggaaacagtaagacagatacgtcacaataaagacaggaagaaacacatctccaatga 475
QY 361 TGTATGATGTTCAAAAAGCCGATGTATCTTCTACAGAGGAGGAGGTATTGATTAAGATC 420
Db 476 tgaatgattgtaaaaaagcagatgtatctctcaagcgagaggtatattgataaagactc 535
QY 421 CTTAGACCGCTTTTACTTCAGGCAATGGATGGTTCTCTATTGTGTGAATCGAGAGC 480
Db 536 cttaagacgccttactctcagcaltgaltggtctcctattgtgtgtaalcgagacg 595
QY 481 AAACATGTATTGTATCAGAAAATGTCCACAAATACCTGCAATTAAGCAAGAGAGCT 540
Db 596 aaacatgtattgtatcagaaaaatgtcacacaatacctgcaataaagcaagagacct 655
QY 541 GGTAAACACAAGTGTTTACAATATCTTACATGAAGAGACAGAAAGATTTTCTTAAGAA 600
Db 656 ggttaacacaagtggttacaatatcttacaatgaagaagacagaagagttcttctaagaa 715
QY 601 TTTACCAAAATCTACAGTTAATGAGTTTCCTGACAAATGACCCCAAGACAAAAG 660
Db 716 ttcccaaaaatctcagttaaatgtggttccgtgcgaataagaccacaagaacaaaag 775
QY 661 CCATCATTTAATTGCGCTATGTTGATGATAAACCACATGATATTCGGAAGACATAA 720
Db 776 ccaatacatltaatgycgtatgtgataagaaacacacacatgatalactcgtgaagacataa 835
QY 721 GCGCAGTCCTGAATGCGCCAGAGATATGAACAATGCAAGTGCTTTGCCCTGTCCAGCC 780
Db 836 cgccagctcgtgaatgycgagagatcgtgaacaaatgcaatgtcgttcgcctgcagcc 895
QY 781 ACGAGCTTGAATGAGAGAAAGGGAAGATTGCAATCTTGTATGATCTGTGTGCGACGCC 840
Db 896 acgagctatgtatgaggaaggaagacttccaatctgtatgtacgtgtgtgtgcacgcg 955
QY 841 CATTACTACAGAGAGAAAGACATTTCATCAAAACCCTGAGAGCTTATTACAGACATGA 900
Db 956 catctactcagggagaaagacatctccaataacccctgagaccttataccaagacatga 1015
QY 901 TCTTTGAGAAAAGTTGTCAATATAGATACAATTTCACTGAGATCCTCATGAGGCTGG 960
Db 1016 tctttcaggaaggtgtgtcaatatagatacaaatctcactgagatctctcatgaagccg 1075
QY 961 CTTTGAACATATTAATCCCAAGGTGTATTCCAGAGATTTTGTAGTCTTAATGATGGCAGTC 1020
Db 1076 ctltgaagataataatcgaaggtgtatcagagatcttcttagtctaaatgtatggtgcaatc 1135
QY 1021 ATGGTCCAGAGAAAGCTCACTATTCAGAAGCTTATCTTAATGAGCCATGAGAAAACCCAGT 1080
Db 1136 atgtgtccagaaaggtcactctcaagaagcttatcttaatgtgtccatgtcagaaccccgat 1195
QY 1081 ATATGATTTCTGTTGGCTGATGGAATATAGTGACTGCACAGACAAAAGCAAACTGT 1140
Db 1196 atatgattctcgtgtgtgtgtgactatagtgactgtcagagacaaaagcaaacctctt 1255
QY 1141 CCGAAATCTGTAAACAAATGATCGACATGGGCTTTGTCTCAACCCACTCTCTTACAGAGA 1200
Db 1256 ccgaataatcgttaacaaatgtatcgacatgtgtctgtccaacccacttctctcagagaga 1315
QY 1201 ACAGAAATGATATAGACAAACCCAAATCTGTTTGACAAGGATTTAGACCACTATGGC 1260
|||||

Db 1316 acagaatgatatagataagccaaacccaaatccctgtgtggaagaaggtatgaaccaactatgyc 1375
QY 1261 TGTATGCAACAGTTTCGTAGCGCGCATGATGATGTGTGCCAAACCAAGGCTTACAGATGCC 1320
Db 1376 tgaatgcacaagcttcgttagtcgagcatgatagtgtcgcacaaccaaagcttacaagatgc 1435
QY 1321 GAGCAGCAGGGGCTTATGGCTTGGCGAGACCTTAGCAACCAAGGCGAGATGATGAGACTTC 1380
Db 1436 gagcagcagggcctatgtgtgtgcagaccctagacccaagggcagatgaatgtgaagctag 1495
QY 1381 GTATGGGGTTTCAGATTAATATAGCTTCAATTAACCCCTGGCGCAGCATCAATACCATC 1440
Db 1496 gtatgtgggttcccgtaacataagcttcatgtacccctgtgcccagatgacataccatc 1555
QY 1441 TTCTTACCAAGAACCAACTATGGGCTCACATGATGATGAGCCGCCACATGGGAGTCTGG 1500
Db 1556 ttccctacagaaacaacactatgtggctcaaatatgtagtcccccaatgtggagttccgtg 1615
QY 1501 TCTTCCCCCAACCGCAGAAATATCATGATTTTCTCTGCTAATGCTGGAGTCCAAAGAT 1560
Db 1616 tcttgccccaacacagcagaatatcatgtattctcctgttaatcgttggagttccaaagat 1675
QY 1561 AGCCATCATGAGTTTCTCCGTTGAGGTTGACATCTCCCATGTCATCTTCGCAA 1620
Db 1676 agcccatcatcagttcttccctgtgcaggtgtgcaactccccaatgcaatctctgtgcaa 1735
QY 1621 TACTGGGAACCAACAGCTTTTCCAGCAGCTCTCTAGTGCCTTCGAAGCCATCAGTGAAG 1680
Db 1736 tactgtggaacacacagcttctcagagatctctcagtgccctgcaagcatcagtgaaag 1795
QY 1681 TGTGGGCACTTCCTTTTATCTACTCTGTCAATCACAGGCCCAATTTGATTAATCTGCC 1740
Db 1796 tgtgtggacttccctcttatactactgtgtcatccacagccccaatgtgaataactctcc 1855
QY 1741 CAATATGAATATTACCAACCAAGTAAAGTAAAGCAATAGGATTCCAAGACTCTCTGG 1800
Db 1856 caataatgatatataccacaacaaatgaagtaagcaatacagaatctccaggtctctgtgg 1915
QY 1801 CTTTATTTGCAACCAAAATCCAGTGAGAGATTCAATGTGTCAATCAATTAACAGATCA 1860
Db 1916 ctlttatgtgcacaaaatccaggtgtgagagttcaaatgtgtcagttcaaatatgacagatca 1975
QY 1861 CCTCAGTACAAGAAATGAAGGAGAGCAGTGTGAGGGGCGACAGAAATCAAGGGGTCC 1920
Db 1976 cctcagtgacaaagaaatgaagagagcaggtgtgtgaggggcagagaaatcaaaaggtgtcc 2035
QY 1921 TTTGAAAGCAAAAGCTCAATAAAAAATTAATCTGAGTAACTTACCTGTTCTTGATGACCG 1980
Db 2036 ttgtgaagcagaaggtgtcaataaaatctatcagttactctacccgttctctgtatgacg 2095
QY 1981 GGGTCATTTCTCTTGACCAACTCCCCCTAGATTCAAGTTGTAAAGATCTTGTATG 2040
Db 2096 ggtgtcatctccctctgtacaaactcccccttagatccaagtgtataaagaatctctgttag 2155
QY 2041 TGTCAACGCCCCCTTGAGTCTCTCTCTTACATCTGTGAGAGAGATATCTTACATCCAA 2100
Db 2156 tgtcaacggccctctgtgagttccctctcaactctgtgaagatlatccttaactcaaa 2215
QY 2101 TATGAGTGGTCACTGTATACAAAGAGAAAGCACCGGATTTTGACAAGTGTGCGAAGATTC 2160
Db 2216 tatgtatgtgtcactgttacaagaagagcaacggaatttltgcaaaagtgtgcgcgaatgtg 2275
QY 2161 GAATTCACACAGCTGAGTAGCCAAAGATTACTGCACAAAGCACTGGGAAAGACACAGCAG 2220
Db 2276 gaattcacacagctgtgtgtgagccaagatatactgcagaagccaactgtgaaagacacaag 2335
QY 2221 TATTAATCTTGTGGGAGACGGAATGTGTCAAGCAGAGACAGCTAAGTCTTAAGAAAGAA 2280
Db 2336 tataactctgt 2395
QY 2281 GGAGAAATATGACACTTCTTAATATACCTGCTGACAGGAGATATCTATGATGACACTTC 2340
Db 2396 gtagaataatgtcactctcttaagataactcgtgtgacagggatgaatccttagtgtacacttc 2455
|||||

QY	2341	TAAGAAGCTACAGCCCCAGATGTGGAGAGATGGACATTAATAATGATGTCAGTGGCCACGACTC	2400
Db	2456	taaaagaaactacagcccccaagtggaagagtggaataaaatgagtcagtgaccagctc	2515
QY	2401	CACCAATTCCTACTCAAGCAAGCAAGAAAGACCCTAAATAATTAAAGACAGAGACAAGTGAAGA	2460
Db	2516	caacatctcctagctcaagtcgaagagaagaaccctaaatlaagaccagagacaagtgaaaga	2575
QY	2461	GGGATCTGGAGACTTGGATTAATCTAGATGCTATTCTTGGTGATCTGACTAGTTCTGACTT	2520
Db	2576	gggactctggagacttggataatcatagatgctatcttctgtgactgactagcttgcactt	2635
QY	2521	TTACATATATTTCCATATCTCTCAATGTGTAGTCATCTGGGGGACTTAACCAACAGTGTTTTCA	2580
Db	2636	ttacaataatccatatactctcaaaatggtagtcactctgggagactaaacaagtgcttca	2695
QY	2581	AGGAACATAATTCCTGGGTTTAAAAAGTTACAGACTGTGTGAGTCTATTTCGTCTCCATAT	2640
Db	2696	aggaactaatctctcgtggtttgaaaaagttcacagctctgtgcagctcattctgcttcacata	2755
QY	2641	TAAACCAGACAGTGTCTCTGGATAGCCCTGTTTCTGTGGCTCAAGTCTTCAGCTAAAAAA	2700
Db	2756	taaccagacagtgctctctgtagagcccttctctgtgctcaagctcccaataaaaaa	2815
QY	2701	TATCATGTGTTTCCCATATTTACCAAGACCCATGTTGGGTGGGAATCCAGATATGAT	2760
Db	2816	tatcagtgcttcccccaatgcttaccaaagaaccaatgctggtggaatccaaagatgcat	2875
QY	2761	GGATAGTACGAAAAATTATGGCTCAAGTATGGGTGGGCCAAACCGAAATGTGACTGTGAC	2820
Db	2876	ggatagtcaggaanaattatgcttcaagtagtggtgggccaacacggaatgtagctgtgac	2935
QY	2821	TCAGACTCTCTTCTCTCAGAGAGACTGGGGCTTACCAACTCAAAAGCCGGCAGAAATGGAACC	2880
Db	2936	tcagactctctctcaagagacttgggtcttaccaaactcaaaagccggagaaatggaacc	2995
QY	2881	TATGCAATTCAAACTCCATCGGAAGACAGAGAGAGATTAATATCTTCTTTACCCTGAGAC	2940
Db	2996	tatgcatccaactccatagygaaagaccagagagagatatatactcttaccaccagacc	3055
QY	2941	TGCACCTGGGTGCTCTATTCCACATTTGCTTCGTCTAATATAGCATACAGGTGCGAG	3000
Db	3056	tgcatctgggtgctctatctcccaactatgctcttgcgtctaaatagataccaggttgcag	3115
QY	3001	ACCAGATTTGCAAGACGACAGACAGATGCTCAATGAGGCTGGTGAATCCCCATGGG	3060
Db	3116	accagatattgcaacagcagcagcagatgcttcaaatgaaagccgtggaatccccaatg	3175
QY	3061	AATGGGGGCTAATCCCTATGAGGCAAGACAGACAGCATTAACAACTGGTTCCTGCGCCGA	3120
Db	3176	aatgggggctaactcctaagtcgaacagcagcagcatctaaaccaacttgytcttcggtccga	3235
QY	3121	TGGCATGTTTTCATGGAACAAGATTTCATGCGACTCAAAATATAGGCTCTTCTTAGAA	3180
Db	3236	tggtcatgtgttccatgynaacaagttctcatgtgcactcaaaatagagcccttctcttagaa	3295
QY	3181	TTCCCTGGATGATCTTGTGGGCGACCTTTCAAACCTGGAAGCGACAGTGAAGCAAGACG	3240
Db	3296	ttccctggatgactcttgttgggccaacttccaaacctggaagcgagatgacgaagaagc	3355
QY	3241	ATTATTTGACACGCTGACACTCTTCTCAGCAACACAGATGCGACAGGCTTGGAAGAAAT	3300
Db	3356	attattggaccagctgcaacaactcttccagcaacaagaatgccaagcgctctgaaagaat	3415
QY	3301	TTACAGAGACTTTTGGGCAATTCCTGTAACCTTCACTACGGAACAGGCAATTAGAGCCCAACA	3360
Db	3416	ttacagagacttttgggcaattccctgtaacttgcatacaggaacagcattagagaccacaaca	3475
QY	3361	GGATGCTTTCCAAAGGCAAGGACAGACAGTAATGATGATGATGATGATGATGATGATGATG	3420
Db	3476	ggatgctttccaaaggccaagacagcagtaatgataatgataatgataatgataatgataatg	3535

```

Db 4616 tgcattctgagcaggact 4633
|||||
RESULT 7
US-60-172-373-13551
; Sequence 13551, Application US/60172373
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Binh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
; FILE REFERENCE: GX-0006 P
; CURRENT APPLICATION NUMBER: US/60/172,373
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 25,772
; SOFTWARE: PERL Program
; SEQ ID NO 13551
; LENGTH: 7956
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO: 445586.3
US-60-172-373-13551

Query Match 98.3%; Score 4421.2; DB 49; Length 7956;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 4483; Conservative 0; Mismatches 8; Indels 27; Gaps 3;

QY 1 GCTGGATGGTGACTCAGAGACCAATMAAATMACTGCTTGACATCCTTTGACTGTT 60
Db 131 gctgtagtgtgacccagagaccataaataacacgtctgacatccttctgactgct 190
QY 61 AGCCAGTGTCTATATATTTCAAGATGAGTGGATTGAGAGAAACCTTGATCCACTGCC 120
Db 191 agccagctgtgtatgtatattcaagaatgagtgagtagagaaaacttgatccactg 250
QY 121 CAGTATTCACGAAAGCAAAATTCATGTGATCTCCAGACAAAGGCTTAACCTGCGAG 180
Db 251 cagtattatccagaaagcaaatctgcattgatactccagagacaagctctcaactg 310
QY 191 TGTGTAAGACGAGACGGAGACAGAGAAAGTAATATATTGAAGATTGGCTGAGCTGAT 240
Db 311 tgtgtaaaaacgagagcagcgagcagaaatataatgaaatgagatcgctgagctg 370
QY 241 ATCTCCCAATCTTACTGATATTGACAATTTCAATGTCAAAACGATTAATGTGCATTTT 300
Db 371 atctcccaactcttagtgatattgacaatttcaatgtcacaacagataaattgtgc 430
QY 301 AAAGAAACAGTAAGACAGATAGTCAATTAAGAGCAAGAGAAACGATTTCCAAATGA 360
Db 431 aaagaaacagtaagacagatagctgcataaagaagcagaagaaacatctccaatga 490
QY 361 TGAGATGTTCAAAAAGCCGATGTATCTTCAACAGGCAAGGAGTATTGATTAAGAGCTC 420
Db 491 tgaatgcttcaaaaagccgagtatctctcaagcgagagtgatctgataaagactc 550
QY 421 CTTAGACCGCTTTTACTTTCAGGCAATTTGATGGTTCCTATTTTGGTGAATCGAGAGC 480
Db 551 cttagacgcgcttactctcagcattgagatggttctcctattctgtgtgaatcgaga 610
QY 481 AAACATTTGATTTGATCAGAAATGTCAACATATCCGCAATTAAGCAAGAGAGCT 540
Db 611 aaacatttatttctatcagaataatgtaacacataactgcaataataagcaagagact 670
QY 541 GGTTAACACAAGTGTTCATATATCTTACATGAGAAAGACAGAAAGATTTTCTTAAGAA 600
Db 671 ggttaacaacaagtgtcttaacaatacttacaagaagaagcagaagaagattttctta 730
QY 601 TTTACCAAAATCTACAGTTAATGAGATTTCCTGAGCAATGAGCCCAAGACAAAAG 660
```

```

Db 731 ttaccaaatactacagtaatgagttccctgcaaaatgagaccacaagaacaaaag 790
QY 661 CCATCATTTAATTCGCGATGTTGATGAAAACACACATCATGATTCGGAAGCATMAA 720
Db 791 ccatcatcttaacttgcgcgtatgtagtgaataacacacacagatattcttgagaataa 850
QY 721 GCCCAGTCCCTGAATGCCCCAGAGATATGAACAATGACAGTGTCTTGCCCTGTCAACC 780
Db 851 gccagctcttgaaatgcgcagagagatgataaacaatgcaatgagcttgccttcagcc 910
QY 781 ACGACTTATGATGAGAGAGAGGGAAGATTTCGCAATCTTGTATGATCTGTGGCAGCCG 840
Db 911 acgactatgagagagaagagagattgcaacttgatgactgtgtagctgtgtgcagccg 970
QY 841 CATTACTACAGGAGAAAGAACTTTCGATGCAAAACCTGAGAGCTTTATACCAGACATGA 900
Db 971 catlactacagagaaagaacatcttccacaaacccctgagagcttattacacagatga 1030
QY 901 TCTTTACAGAAAGTTGTCAATATATAGATACAAATTCACAGAGATCCTCATGAGGCTGG 960
Db 1031 tctttcaggaagagctgtgtcaataatagatacaaaatccactgagatcttcatagagcc 1090
QY 961 CTTTGAAGATATATCCGAAAGTGTATTCAGAGATTTTGTACTTAATGATGGGCACTC 1020
Db 1091 ctltgaagatataatccgaagtgatctcagaagatttlttagtctaaatgtagcagctc 1150
QY 1021 ATGGTCCCAAGACGTCCTATCAAGAAAGCTTATGTTAATGGCCATGCGAAGAACCCCGT 1080
Db 1151 atggtcccaagaacgctccatacaagaagcttataatgagccatgagaaaccccg 1210
QY 1081 ATATGATTTCTGTTGCTGATGATGAACTATAGTGACTGCACAGACAAACCAACTCTT 1140
Db 1211 atatcgatctcgtgtgcgtatgagactatagtgactgacagaaacaaacactctt 1270
QY 1141 CCGAAATCTCTGAACAAATGATCGACATGGCTTTGTCTCAACCCCTTCTTCAAGAGA 1200
Db 1271 ccgaatctctgtaacaaatgatacgaatgagctgtgtcccaacccactctctcaga 1330
QY 1201 ACAGAAATGATATAGACCAAAACCCAAATCGTTTGGACAAAGGATTTAGACCACTATG 1260
Db 1331 acagaaatgatatagataccaaacccaaatccgtgtgaaagaagatagaaacccata 1390
QY 1261 TGGATGCAACAGTTGCGTAGCGGCGCATGATATGTGCGCAAAACCAAGCTTACAGATGCC 1320
Db 1391 tgaatgcaacagttcgttagcgtagcgatgatagtgcgcaaaacaaagcttacaagat 1450
QY 1321 GAGCAGCAGGCGCTTATGGCTTGCGAGACCCCTTAGCACACAGGCGAGATGATGGAGCTAG 1380
Db 1451 gagcagcagggctatgagcttgcgcagaccctagcaccaaggcagatgagtgagctag 1510
QY 1381 GTATGGGGTTTCCATTAACATAGCTTATTTGACCCCTGGCGCAGGCAATGCAATCCATC 1440
Db 1511 gtatgggggttccagtaacataagcttcaatgtgacccctgycgcagagcagatacc 1570
QY 1441 TTCTCAACGAAACAACAACATATGGGCTCAACATGAGTAGAGCCCCCAATGGAGATCTCGG 1500
Db 1571 ttctcaacgaagaacaactatgagctcaacaatgagtagcccccaatgagagctc 1630
QY 1501 TCTTCCCCCAACGACGAATATCATGATTTTCTCTGTAATGCTGGAGATCCAAAGAT 1560
Db 1631 tcttcccccaacccagcagaatatacagattctctctctgtatctgtgagatccaaag 1690
QY 1561 AGCCATCATCATGATTTTCTCCGTTGAGAGGTTGACATCTCCCAATGGCATCTTTCGCAA 1620
Db 1691 agcccaatacatgatttctctcgtgtgcagagtgtagcaactctcccatgagatctctg 1750
QY 1621 TACTGGGAACCAACAGCTTTTCCAGCAGCTCTCTCAGTCCCTCGCAAGCCATCAGTGAAG 1680
Db 1751 tactgggaacacagcttctccagagctctctcagtgccctcgaaagcatalcagtgaa 1810
QY 1681 TGTGGGGAATTCCCTTTATCTACTCTGTGCATCACAGGCCCCAAATTTGATTAATCTGCC 1740
```

Dd	1811	tgybgyacttcccttctatctctctgtcctacccaggccccaattgatacttcc	1870
Qy	1741	CAATATCAATTAATTAACCCAAACAGTAAGTAAGCAATCAGATTCCTCTGG	1800
Dd	1871	caataigaatatttaccacccaagtaangtaagcaatcaagttccaaagtctctggy	1930
Qy	1801	CTTTTATTCGACCAAAATTCAGTGGACAGTTCATGTGTCAGTCAATATCAGAGATCA	1860
Dd	1931	ctttattcgacaacaatactcagtggaagattcctaattgtcagtcacaataagcagaatca	1990
Qy	1861	CCCTAGGACAAAAGAAATAGGAGAGAGAGTGTGGGGGGCGAGATCAAGAGGGTCC	1920
Dd	1991	ccctcagtgacaagaagaagtaaggagacagtgctgaaggggcagaagaatcaaaagggtcc	2050
Qy	1921	TTTGTGAAGCAAAAGTGCATATAAAATAATCTGAGAGTACTTACTCTTCTTGATGACCG	1980
Dd	2051	tttgaagaagcaaaagtgatataaanaattactcagttacttactcttctctctgtatgacg	2110
Qy	1981	GGGTGATTCCTCCTTGACCAATCTCCCCCTAGATTCAAGTGTAAAGATCTTCTGTTAG	2040
Dd	2111	gggtcattctcctcttgacacactccccctagatcgaattgtlaagaatctctgttag	2170
Qy	2041	TGTGACAGGCCCTCGTGGAGTCTCTCTCATCTGATGGAGAGATCTCTGACATCCAA	2100
Dd	2171	tgtaaccagcccccttgagtgctccctccctctaactctggaaggatctcctctaactcaa	2230
Qy	2101	TATCATATGGGCTACTGTTACAAAGAGAAGACCGAGATTGTGCAAGTTGCTGCAATATGC	2160
Dd	2231	tatgcatgggttcactgtltaaaagaagacacggatattgtcacaagttgtctgcagaatg	2290
Qy	2161	GAATTCAACAGCTGAGGTAGCCCAAGATTACTGCAACAGCCACTGGGAAAAGACACCAGAG	2220
Dd	2291	gaattcaaccagctgagtgtagccaagttactctgaagaagcactggygaagaacacagag	2350
Qy	2221	TATTAATCTTGTGTGGGACGGAATGTGTGCAAGCAGAGACGTAATCTCTAAGAGAA	2280
Dd	2351	tataactctctgtggygcgaagaattgtltaacagcagagcagctaaagttccctaagaaga	2410
Qy	2281	GGACAAATATATGCACTTGTGATACCGGCGGACAGAGATGATCCGAGATGCACTGTC	2340
Dd	2411	ggagaaataagcaactctctatgaataccctgcgacagggatgctccagtgagacactcc	2470
Qy	2341	TAAAGAACTACAGCCCCAAGTGGAGAGAGTGCATTAATAAAGAGTCAAGTGCAACGCTC	2400
Dd	2471	taagaactactacagcccacaagtgggaaggtgtgacaaataaatgagtgactgacacagtc	2530
Qy	2401	CACCATTCCTAGTCAAGTCAAGAGAGAAGACCCCTAAATTAAGACAGACAGATGAAGA	2460
Dd	2531	caccattctctagctcaagtcaagaagagaagaaccctaataatlaagacagagacagtgaaga	2590
Qy	2461	GGGATTCGGAGACTTGGAATATCATGATGTATTCTGGTGATCTGACAGTTCGACTT	2520
Dd	2591	gggactcgggaacttggataaactcagtgctatctcttggatctcgacactagttcgtactt	2650
Qy	2521	TTACAAATTAATTCATATCTCAATAGGTATGATCATCTGGGAGATTAAGCAACAGTGTTC	2580
Dd	2651	ttacaataattccatctaccctccaatgtagtactcggggactaagcaaacagtgtttca	2710
Qy	2581	AGGAATCAATTTCTGTGGGTTTGAAAAGTTCACAGTGTGACAGTCAATTTCTGCTCCATA	2640
Dd	2711	aggaaactaatctctcgttggttgaagaagtacacagctctgtgcagctctatctgcctcata	2770
Qy	2641	TAAACGAGCACTGTCTGTGGATGAGCCCTGTTCTGTGGCTCAAGCTCCAGTAAATAA	2700
Dd	2771	taacggagcagtgctctcgtgatagcccgtttctgtgtgtcctaagctcccacgtaaaaaa	2830
Qy	2701	TATCAGTGTCTTTCCCACTGTGTACCAAGCAAACCAATGTTGGTGGGATCCAAAGATAT	2760
Dd	2831	tatcagtgcttcccatgtltaacaaagaaccatggttgggttggaatccaagaatgat	2890
Qy	2761	GGATATGACGAAAAATTATGGCTCAAGATGAGTGGTGGCCAAACCGAATGTGACTGTGAC	2820
Dd	2891	ggatagtgtaaggaanaattatgtctcaagtaaggttgggtggcacaacgcgaatgtgtaactgtac	2950

OY	2821	TCCACCTCCTTCTCAGGAGACTGGGCTTTACCAACTCAANGCGGCAGATGAAC	2880
Db	2951	tccagcttccttcccagagactggggtttaccaaactcaaggccggcagaatggaac	3010
OY	2881	TATGATTCAAACCTCCATGGGAGACCAGAGAGATTTAATTCTCTTAACCAAGCC	2940
Db	3011	tataattccaactcccatgggaagaccagagatatataacttcttaaccaagcc	3070
OY	2941	TGCACGTGGTGCGTCATWTCCCACATTGGCTCTTGCGTTAATACATACAGGTGCGAG	3000
Db	3071	tgcactygytgctcttatcccaattgctcctctcgylctaataagcalaccagytgagag	3130
OY	3001	ACGAGTATTGCACAAGCAGCAGCATGCTTCAATGA-----GCCGTGTGAAT	3051
Db	3131	acsagiatltgacaagacagcagcagatgtcttaaatgactccccagggcttygaaat	3190
OY	3052	CCCCATGGGAATGGGGCTTAATCCCTAATGGCCAAGAGAGATTTAACCACTGGCTTC	3111
Db	3191	ccccatlygnaatlygggtctaacctccctatgyccaagcagcagcatlaaccaactlygtlc	3250
OY	3112	CTGGCCGATGCAWTTGTTCACATNGAACAGTTTTCTCATGGCACTCAAAAATAGGCCCT	3171
Db	3251	ctggcccgatlygcatlygtlccatlygaacaagtlltcatlygcaactcaaatagacctct	3310
OY	3172	TCTTAGGAATCCCTGGATGATCTTGTTHGGGCCACCTTCAACCTGGAAGGCCAGATGA	3231
Db	3311	tcttaggaattccctgtagatctgtfityggccaacttccaacctggaagsgcagagatga	3370
OY	3232	GGAAGAGCAATTANTGGACCAAGCTGCACACTCTTCTCACCAACACAGATGCACAGCCT	3291
Db	3371	cgaaagacatlatlygaccagctgycacactcttccagcaacaagatlycacacagacct	3430
OY	3292	GGAGAAATTGACGAGACTTTGGGCAATTOCTGACATCTGTCAATCAAGGAGACGGCAATTGA	3351
Db	3431	ggaagnaattlygacagagcttlyggcatctccgaaactlgtcaalcasggygacgatlaga	3490
OY	3352	GCCCAACAGAGATCTTCTCCAAAGGCCAAGAAGCAGCAGTAATGATGATCAGAAAGCAGG	3411
Db	3491	gcccaaaagatlygtcttccaaagycceagaagcagatgaatlygatcagaagagcag	3550
OY	3412	ATTATATGACGACATACCAGACAGGGGCTCCCAATGACAGGAGGCTTTCATCTTCA	3471
Db	3551	attatatgacagcatataccocagacaggggtcccaatlygaagagaggttcatcttica	3610
OY	3472	GGGACATCACCATCTTTTAACCTATGATGATCAAGATGAAMCAGCAAGCAATTTCC	3531
Db	3611	gggaacaatcacacacttctaacttaagatagaatcagaatgaaacagcaagaaatttcc	3670
OY	3532	TCTCCAAAGATGACCCACGAGCAATCATGAGAACCCCGGACAACACCCCCCAACA	3591
Db	3671	tctccaaggaatlygacccaacagacatcatagaaaccccggaacaacaacccccaaaga	3730
OY	3592	ACTTAGAATGCAAGCTTCAGCAGAGAGGCTGACGGGCCAGAGTTTGAATCAGAGCCGACA	3651
Db	3731	acttagaatygagttcagcagagaggtctgcaagcgaagttttaagatcagaagcgaca	3790
OY	3652	GGCACTTAATTTGAAAAATGAAAAACCTACTGCTGGTGGCTGGCGGTGATAGGCTAT	3711
Db	3791	ggcaactttaatlygaaaaatlygaaaacccctaactgtctgtygtctgcgtyatlyagacctat	3850
OY	3712	GATGCAGC-----CCGACGAGGGTTTTCTTAATGCMCAAAATGTTGGCCACAG	3759
Db	3851	gatgcagccccaggtlygagctcccaagcaggttlltctaagtctcaaatlygttcgccaag	3910
OY	3760	CAGCAGAGAGCTGCTAAGTATCATCTTCCGACAACAGAGGTTGGCTATGATGATGACACA	3819
Db	3911	cagcagagagctgtctaagtcatcaacttcgacaacaagaggytgygtatgatgatgcagca	3970
OY	3820	GCAGCAACAGCAGCAGCAGCAGCAGCAG-----CAGCAACAGCAACAGCAACAGCA	3873
Db	3971	gcagcagacagcagcagcagcagcagcagcagcaacagcaacagcaacagcagcaacagca	4030

QY	3874	GCACAGCAGCAAAACCCAGGCTTACGCCCACTCTCTAATGTGACTGCTTCCCCAGCAAT	3933
Db	4031	gaaacagcagcaaaacccaagcctccagccacctccctaaatgtacatgctctcccccagcat	4090
QY	3934	GGATGGGCTTTTGGCAGGAGCAACATAGCCACAAAGCTCTCCGCAACAGCTTTCATATCA	3993
Db	4091	ggatgggctcttttggcagggcccaatgccaagaagctctctccgcaaatccatata	4150
QY	3994	ACCAATATTATGGAATGGAGCAACAAACAGATCCAGGCTTTGGTCAGTGTCTAGTCTTCC	4053
Db	4151	accaaatatggaatggagcaaaccaacagatccagccttggctagctctcc	4210
QY	4054	CATGTGAATGATGTGCTCAAGAATGGGTGCTCTCCGAAATCCCATATGCAACACCCGCA	4113
Db	4211	catgtgaatgatatgtctcaagaatgggtctctccgaaatcccatatgcaaacccgca	4270
QY	4114	GGCTGCATCATCTATACATGCTCCATCAAAATGAAGGCTGGCCATCGAAATTTGGCCAG	4173
Db	4271	ggctgcatacctatcacgtctccagaatgaaggctggccaatcggaatcttggccag	4330
QY	4174	GAACAGCTCTCTTTTCCACAGACAGATTTTCCACACAGGGAAATCCGACGTGTATGAT	4233
Db	4331	gaacagctctcttccacagcagatgttggccacagggaatcctcgagtgatagat	4390
QY	4234	GGTGCACATGAATGGCAGCAGTGGTGCATGTGGACAGATGAACATACCCATGCCCAT	4293
Db	4391	ggtgccacatgaatggcagcagtggtgcattggtggacagatgaacataccatggccat	4450
QY	4294	GCTCGCATGCTTATGAGGTGCTGATCAGAAATACCTGACATCTCTGACACAGACGTC	4353
Db	4451	gctcggcatgcttatgggtgctgattcagaaatacctgacatctctgacacagacgcttc	4510
QY	4354	TTAAGGAACACACTGTACAATGACACTGCTAGGATTATTGGGAAGAAATCATTTGTTTC	4413
Db	4511	ttaaggaacacactgtacaatgacactgcttaggattattgggaagaaatcatattgttc	4570
QY	4414	CAGGCATCATCTTGGAAAGAACACCACTTTGAGACTTCATCAGAGGATTTTAAGTGA	4473
Db	4571	caggcatcatcttggaaagaaacacactttgagacttcacatcagaagttatttaagtga	4630
QY	4474	TGTCATTTTGACGAGCAAT	4491
Db	4631	tgctcatlttgagcaggaact	4648
RESULT	8		
	US-60-209-009-198		
	: Sequence 198, Application US/60209009		
	: GENERAL INFORMATION:		
	: APPLICANT: Paris, Mary		
	: APPLICANT: Pearson, Cecelia I.		
	: TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER		
	: FILE REFERENCE: PA-0027 P		
	: CURRENT APPLICATION NUMBER: US/60/209,009		
	: CURRENT FILING DATE: 2000-06-01		
	: NUMBER OF SEQ ID NOS: 501		
	: SOFTWARE: PERL Program		
	: SEQ ID NO 198		
	: LENGTH: 7116		
	: TYPE: DNA		
	: ORGANISM: Homo sapiens		
	: FEATURE:		
	: NAME/KEY: misc_feature		
	: OTHER INFORMATION: Incyte ID NO: 1094199.1		
	: NAME/KEY: unsure		
	: LOCATION: 3941-3993, 5899-5939, 6991		
	: OTHER INFORMATION: a, t, c, g, or other		
	US-60-209-009-198		
Query Match	97.7%;	Score 4393.8;	DB 52; Length 7116;
Best Local Similarity	98.4%;	Pred. No. 0;	
Matches 4437: Conservative	0;	Mismatches	54; Indels 19; Gaps 2;

QY	1	GCTGGATGCTGGACTCAGAGACCAATAAAAATAAAGCTTTGAACATCCCTTGACTGGTT	60
DB	116	gctgtagctgtagactccagagcccaataaaataacatgctctgaacatcccttgactgctt	175
QY	61	AGCCAGTTGGCTGANGTATATTTCAAGATGAGTGGATTAGGAGAAAAGTTGGATCCACTGGC	120
DB	176	agccagcttgcgtagtatactcaagaatgagctgagcttagagaaaacttggatccacttgc	235
QY	121	CAGTATTTACAGAAAACGCAAAATGGCATGTGATATCCAGCAAGGCTTAACTACCTCAG	180
DB	236	cagtgatcccaagaaaaagcaaatggccatgtagtaactcccgagcaagtgcttaactcgcag	295
QY	181	TGGTAAAAAACGGAGACGGAGCAGAGAAAGTAATATATTTGAAGAAATTTGGCTGACTGAT	240
DB	296	tggtagaaaaacggagacggagcgaggaagtaatatatactgaagaaatctgcttagagctgt	355
QY	241	ATCTCCCAATCTTAGTGATATTTGCAATTTTCATGTCAAAACGATTAATGTGCCATTTT	300
DB	356	atctcccaatcttagtgatatctgcaaatccaaatgtaaacacagaaataatgtgcgattct	415
QY	301	AAAGAAACAGTAAGACAGATACCTCAATTAATAAAGACCAAGAAAACATATTTCCAAATC	360
DB	416	aaagaaaacagtaagacagataagctcaataaagaagcaagaaaactatctccaaatga	475
QY	361	TGATGATGTTCAAAAAGCCGATGATCTTTACAGGGCAGGAGTTATTTGATTAAGACTC	420
DB	476	tgatgattgtcaaaaagccgattgattctctctacagggcgagggtattgtataagagctc	535
QY	421	CTTAGACCCGCTTTTACTTTCAGCATTTGGATGGTTTCTTATTTGTGTGTAATTCAGAGCC	480
DB	536	cttagaccgcttcttaactcagcatttggatgtgttctctatttctgtgtaactgagacgg	595
QY	481	AAACATTTATTTTGTATTCAGAAAATGTCAACAATACCTCAATATTAACAAGAGACCT	540
DB	596	aaacattgatttgaatacagaanaatgtccacaataactctgcaataaagcaagaaggacct	655
QY	541	GGTATACACAAGTGTTAACAATATCTTACATGAAGAAGACAGAAGATTTTCTTAAGAA	600
DB	656	ggtatcaccaagtgcttacaatacttacaatgaagaagacaagaagattctcttaaga	715
QY	601	TTTACCAAAATCTACAGTATATGAGTTTCTGTGACAAATGAGCCCCCAAGCAAAAAG	660
DB	716	tttaccaaaactcacagttaaatggagtttctcttgcgcaaaatggagaccaaaagaag	775
QY	661	CCATCATTTTAAATGGCTATGTTGATGAAGAAACCCCATGATATTTCTGGAAGACATAA	720
DB	776	ccatcatattaatgtccgtagtattgtgaagaaaaccacaagatattctgtgaagaactaa	835
QY	721	CGCCAGTCTGAAATGCGCCAGAGATATGAACATGCACTGCTTGGCCCTGTCTCAGCC	780
DB	836	cgcagctctgaanaatgcgcccagagataatgaaaaacatgacgtgtcttgcctgtctcagcc	895
QY	781	ACGAGCTATGATGGAGAAAGGGGAAGATTTTGCATCTTGTATGATCTGTGTGGACGCC	840
DB	896	acgagctatgtatggaggaagggagaaattgcgaactcttgatgatactgtgtgcgcgcg	955
QY	841	CATTACTACAGAGAAAGAAATTTTCATCAATAAACCCTGAGAGCTTTAATACAGACATCA	900
DB	956	catlactacaggaagaagaacaattccacccaacccctggagagcttattaccagaatga	1015
QY	901	TCTTTACAGAAAGTGTGCATATATAGATTAACAATTTCACTAGATCTCTCATAGAGCCTGG	960
DB	1016	tctttcaggaaggtgtgtcaatatagataaaaattcaactcgtgagatccttccaatgaagctgtg	1075
QY	961	CTTTGAGATATTAATCCGAAGCTGTATTCAGAGATTTTATAGTCTAATGATGGCGACTC	1020
DB	1076	ctttgagataataatccgaagtgattatcagagattctttagctaaatgaaatgagtcgagctc	1135
QY	1021	ATGGTCCCAAGAAAGCTACATATCAAGAGCTTATTTAATGGCATGCGAAAACCCCACT	1080
DB	1136	atgggtcccaagaagctacatactcaagaagcttctctaaatggcgaatgagaaaaccagct	1195

QY 1081 ATATGATTTCTGTTGGCTGATGAACTATATGACTGCACAGACAAACAAACTCTT 1140
|||||
Db 1196 atatgattctctgttgctgtaigaaactatagtgactgcacagacaaaagcaactctt 1255
QY 1141 CCGAAATTCCTGTACAAATGATGACATGGCTTTGTCTCAACCCACTTCTTACAGAGA 1200
|||||
Db 1256 ccgaatccctgttaacaatgatcatgactgtcttctgtcaaccacttctctcagagaa 1315
QY 1201 ACGAATGATATAGACCAACCAAACTCTGTTGGACAAGGATTAAGACACTTATGGC 1260
|||||
Db 1316 acgaatgatacagacaaacccaaatccctgttgacaaggatctagaccactatgac 1375
QY 1261 TGGATCAACACTTCCGTAGAGCGCATGATGATGCGCAACCAAGGCTTACAGATGC 1320
|||||
Db 1376 tggatgcaacagtctgttgagcggtcatgtagtgcgcaacaaagctctaaagatgac 1435
QY 1321 GACGACAGGCGCTATGCTTGGACAGCCCTAGCCACACAGGCGAGATGATGAGCTAG 1380
|||||
Db 1436 gagcagcagggcctatgctgtgacagaccctagcacacagggcagatgagtgagctag 1495
QY 1381 GTATGGGGGTCCAGTAACATACCTTATGACCCCTGGGCGACGATGCAATCACCATC 1440
|||||
Db 1496 gtaagggtgtccagtaacaaactatgactatgacccctggcgagcatgcaatcacacatc 1555
QY 1441 TTGCTACAGAACAACTATGGGCTCAACATGATGAGAGCCCGCATGAGAGTCCGTG 1500
|||||
Db 1556 ttctctaccgaacaaactatgactatgacatgagtagccccacacagggagctccg 1615
QY 1501 TCTTGGCCCAACACAGACAGATATCATGATTTCTCTGTAATGTTGGAGTCCAAAGAT 1560
|||||
Db 1616 tcttgcccaaacagcagatatcatgattctctcgttaactcgtggaggtccaaagat 1675
QY 1561 AGCCTACATCATGTTTCTCTGTTGACAGTGTGACACTGCCATGSCATCTTGTGCAA 1620
|||||
Db 1676 agcctacacagttctctcgttgcaaggtgacactccacatgycatctctctgcaaa 1735
QY 1621 TACTGGGAACCAACAGCTTTTCCAGACGCTCTCAGTGCCTTCAAGCCCTCAGTGAAG 1680
|||||
Db 1736 tactgggaacacaaagctttctccagcagctctcctcagtgccctgcaagccatcagtgaaag 1795
QY 1681 TGTGGGACTTCCCTTTTATCTACTCTGTCATCAGCAGGCCCAATTGGATACTCTCC 1740
|||||
Db 1796 tgtgggacttccctttactactcctgcatcacagggcccaattgtaatactctcc 1855
QY 1741 CATTATGAATATTAACCAACCAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1800
|||||
Db 1856 caatagaataatacccaacaaagtaagtaagcaatcagatctcagaagctccctcgg 1915
QY 1801 CTTTATTTGGACCAAAATTCAGTGAAGTGAATGTCAGTCAAAATGACAGATCA 1860
|||||
Db 1916 ctttatctgacaaaataatccagtgagaggtlcaatggtgcaatcaatagcagagataca 1975
QY 1861 CCTCAGTACAAAGAAAGTAAGAGAGACAGTGTGAAGGGGACAGAAATCAAAAGGGTCC 1920
|||||
Db 1976 cctcagtgaaagaagaagtaagagagcagtgltgagggggcagagaatcaaaagggtcc 2035
QY 1921 TTTGGAAAGCAAGGCTATTAATAAATAATTAAGTACTTACTTCTTCTATGACCG 1980
|||||
Db 2036 ttgggaaagaagaagtaataaataatcagtgacttaactcgttctctctatgacg 2095
QY 1981 GGGTCAATTCCTCTTACCAACTCCCGCTAGATTAAGTGAAGTAAGTGAAGTAAGTGAAG 2040
|||||
Db 2096 gggtaattctctctctgtaacaaactccccctcaggtlcaaggtgtaaaagaatctctgttag 2155
QY 2041 TGTACACAGCCCTCTGAGACTCTCTCTCTATCATCTGAGAGATATCTCTACATCAA 2100
|||||
Db 2156 tgtcacacagccctctgaggtctcctcctcatcatcgtggaggttctcctcatatccaa 2215
QY 2101 TATGACATGGGTCACTGTTACAAGAAAGACCCGATTTTGGACAAGTGTGTGAGATG 2160
|||||
Db 2216 taagcatgggtcactgtltaaaagagaagcagcagatcttgacaaagltgttgagaatgg 2275
QY 2161 GAATTACACAGCTGAGGTAGCCAAAGATTACTGCACAAGCCTAGTGGAAAGACACGACAG 2220

|||||
Db 2276 gaattcacagctgaggtagccaagattactgcagaagccactgaggaaagacacagcag 2335
QY 2221 TATAACTTCTTGGGGAGCGGAAATGTTTCAAGCAGAGACACTAAGTCTTAAGAAAGA 2280
|||||
Db 2336 tataacttcttggtggagcggaaatgtgtcaagcagagcagctaaagtlccttaagaaga 2395
QY 2281 GGAGATTAATGCACTTCTTATGATACCGTGGAGCAGGAGATGATCTAGTATGATCATCTC 2340
|||||
Db 2396 ggaagaataagcaatctcttagataccctgcgcagaggaatgataccatgatagacatc 2455
QY 2341 TAAAGAACTACAGCCCAAGTGAAGAGAGTGAACAAATTAATGATGATGACACGACTC 2400
|||||
Db 2456 taagaactacagccccagtggaagagatggtgataataatgagtcagcagccagctc 2515
QY 2401 CACCATTCTGACTCAAGTCAAGAGAAAGCCCTAAATTAAGACAGAGCAAGTAAGA 2460
|||||
Db 2516 caccattctcagctcgaatcaagtaagaagaagccctaaatlaagacagagcaagtgaga 2575
QY 2461 GGGATCTGGAGACTTGGATATCTAGATGCTATTTGTTGGATCTGATCTGATCTTCTACTT 2520
|||||
Db 2576 gggatctggagacttgataatctagatgctatcttctgtgactgactgactgact 2635
QY 2521 TTACAATTAATTCATATCCCTCAATGATGATGATGAGGAGCTAAGCAACAGGCTGTTCA 2580
|||||
Db 2636 ttacaataatccatatctcccaatgagtagtcatctggggactcaagcaagtglttca 2695
QY 2581 AGGAACATAATCTCTGGGTTTGAAGAAAGTTCACAGTCTGTGACGCTATTCCTCCATA 2640
|||||
Db 2696 aggaactaatctctctggtgttgaaagttcacagctctgacgtctgcatctatctgctccata 2755
QY 2641 TAACCGAGAGTGTCTCTGATAGCCCTGTTTGTGTTGGTCAAGTCTCCATTAAGAAA 2700
|||||
Db 2756 taaccgagcagtgctctctgataagccctgttctgtgtgcaagctccacagtaaaaaa 2815
QY 2701 TATCAGTGTCTTCCCATGTTTACCAAGCAACCCAGTTTGGTGGGAATTCAGAAATGAT 2760
|||||
Db 2816 tatcagtgcttccctccatgcttaccaaagcaaccatggttgggtggaatccaaagatgat 2875
QY 2761 GGATPAGTACAGAAATTAATGAGTCAAGTATGAGTGGGCCAAACCGAAATGTGACTGTAC 2820
|||||
Db 2876 ggaatgcaagaaatatgtgtccaaatgagtgatgggtggccaaacggaatgtgactgtgac 2935
QY 2821 TCAGACTCTTCTCCAGGAACTGGGGCTTACCAACTCAAGCCGCGCAATGGAAC 2880
|||||
Db 2936 tcaagactcctctccagagacggtgggttaccaaactcaaaagccggagaaatggaac 2995
QY 2881 TATGAATTCAACTCCATGAGGAAGACAGAGAGAGATTAATTAATCTTATACCAAGACC 2940
|||||
Db 2996 tatgaattcaactccatggaagacagagagagatataaactcttcttaccagacc 3055
QY 2941 TGCACGTGGTGTCTATTTCCACATTTGCTTTCGCTTAATAGCATACAGGTGCGAG 3000
|||||
Db 3056 tgcacgtgggtgtctatcttccacatgctcctctggtctaaatagatlaacaggtgag 3115
QY 3001 ACCAGTATTCACACAGCAGAGAGATGCTTCAAAATGAGGCTGTGGAATTCCTCATGGG 3060
|||||
Db 3116 accagatltgcaacagcagcagagatgcttcaaatgaggtccgtgtgaaatcccatgg 3175
QY 3061 AATGGGGCTAATCCCTATGAGCAACAGAGCATGTAACCAACTGGTCTCGGCCGA 3120
|||||
Db 3176 aatggggctaaatccctatgagccaaagcagagacatlaacaaactggtgttccgcccga 3235
QY 3121 TGGCATGTTGTCCATGAAACAAGTTTCTCATGGCTCAAAATAGGCTCTTCTTAGGA 3180
|||||
Db 3236 tggcatggtgtccatgaaacaagttctcattgacactcaaaatagaccctctcttaggaa 3295
QY 3181 TTCCCTGAGATGATCTTGTGGCCACTTTCACACTGGAAGGCCAAGTGCAGCAAGAG 3240
|||||
Db 3296 ttcctctgagatctgtgtggccacttccaaactggaagcagagtgagaaagagc 3355
QY 3241 AATATTGACACAGCTGACACTCTTCTACAGCAACAGATGACAGGCTGGAAGAAAT 3300
|||||

D	b	3356	a t u c t t g a c c a g c t g c a c e c t t t c c a g c a a c a c a g a t g c c a a g c c t y g a g a a t	3415
OY		3301	TGACAGAGCTTTGGGCATTCCGTGAACCTTATCATCAGGACAGGCCATTAGACCACAAC	3360
D	b	3416	t g a c a g a g c t t t g g g c a t t c t c t g a c t t g t l c a a t c a a g g a c a g g a t t a g c c c a a c a	3475
OY		3381	G G A T G C T T T C C A A G G C C A A A C A G A C A G T A T G A T G A T C A G A A G G C A G A T T A T A G G	3420
D	b	3476	g g a t g c t t c c c a a g g c c a a g a a g c a g c a g t a a t g a t g a t g a t c a g a a g a g a t a t a t g	3535
OY		3421	A C A C A C T A P C C A G A C A C A G G G G C C T C A A T G C A A G A G G C T T T C A N T M C M A G G G C A T C	3480
D	b	3536	a c a g a c a t a c c c a g c a a a g g g c c t c c a a t g c a a g a g g c t t c a t c t c a g g a c a t c	3595
OY		3481	A C C A T C T T T T A A C T A T G A T G A T G A T G A A C A C G A A G G C A A T T T T C C T C T C A A G G	3540
D	b	3596	a c c a t c t t t a a c t a t g t a t g n a t a g n a t g a a c a a g a a g g a a t t t c t c t c c a a g	3655
OY		3541	A A T G C A C C C A G A G C C A C A C A T C A T G A G A C C C G G A A A C A C C C C A A G A C T T A G A T	3600
D	b	3656	a a t g a c c c a c a g a g c c a a c a t c a t g a a c c c g a a c a a c a c c c a a g c a a c t a g a t	3715
OY		3601	G C A C T T T C A G A G A G G C T G C A G G G C C A G A G T T T T G A T C A G A C C G A C G A G G C A C T T T G A	3660
D	b	3716	g c a g c t t c a a g a a g g g c t c g a a g g c a a g c a g t t t t t g a a t c a g a c g a c a a g c a c t t g a	3775
OY		3661	A T T G A A A T G G A A A A C C T A C T G C T G T G G T G C T G G G T A G A G A G A G C C T A T G A T G A C -	3719
D	b	3776	a t t g a a a t t g g a a a a c c c t a c t i g c t g t g t g t g t g g t g a t g a g g c c a t t g a t c a g c	3835
OY		3720	- - - - - C C A G A G G G T T T T C T T A T G C T C A A T G T G C C C A A C G A C A G A C A G A	3786
D	b	3836	c a a g t g a g c t c c c a c a g a a g t t t c t t c a a t g t c a a t g t c g c c a a g c a a g a g a	3895
OY		3769	G C T C T A A G T A C A T A C T T C C G A C A C A G A G G G G C T A T G A T A G A T G A C A G A C A C A C A	3828
D	b	3886	g c t c t t a a g t a c a t c a c t c c g a a c a a g a g g t g t g c t a t g a t h n n n n n n n n n n n	3955
OY		3829	G C A C A C A C A G A C A G A C A G A C A G C A G C A A C A G C - - - - - A A C A G C A A C A G C A A C A G C	3881
D	b	3956	n n	4015
OY		3882	A G C A A C C C A G G C C T T C A G C C A C A C T C T A A T G T A C T G C T T C C C C A G A C A T G A T G G G C	3941
D	b	4016	a g c a a c c a a g c c t t c a g c c a c c t c c a b t g t a g c t t c c c c a g a t g a t g a t g g c	4075
OY		3942	T T T T G G A G A C C C A C A A T G C C A A G C T C T C C G A C A G T T T C A T A T C A C C A A T T	4001
D	b	4076	f t t t g g a a g a c c c a a t b g c a a a g c t c t c c g a a c a a g t t c a t a t c a a c c a a t t	4135
OY		4002	A T T G A A T G G G A A C A A C A C A G A T C A G A C C T T T T G T G A T G T C T A G T C T C C A A T G C A A	4061
D	b	4136	a t g a a t g y g g a a c a a c c a a c c a g a t c c a g c t t t g t c o a t g t c t a g t c c t c c a a t g a a	4195
OY		4062	T G A T G T G T C A A G A A T G G T C C C T C C A G A A T T C C A T G A T G C A A C A C C G A G G C T G A T	4121
D	b	4196	t g a t g t c t a a g a a t g y g t c c t c c a g a a t c c a t g a t g a a c a c c g a c g t g t a t	4255
OY		4122	C C A T C T A C A G T C C T C A G A A T A A A G G C T G G C A T C A G A A A T T T T G G C A G A A C A C T	4181
D	b	4256	c c a t c t a t c a g t c t c c a g a a t g a a g y g t g t g c a t c a g a a a t t t t g c a g a a c a g c t	4315
OY		4182	C C T T T T C C C A G A C A G A T T T G C C C A C A G G G A A T C T C A G T A G T A T G T A T G T G G C A	4241
D	b	4316	c c t t t c c a g c a g t t t g c c a c c a a g g a a t c c t c g a t g a t a t g t a t g t a t g t g c a	4375
OY		4242	T G A A T G C A C A G T G T C A C A T G G G A C A G A T G A A C T G A A C C C A T G C C A T G T C T G C A	4301
D	b	4376	t g a a t g c a g c a g t g r g t c a a t g y g a c a t g a a c t g a a c c a t g c c a t g t c t g c a	4435
OY		4302	T G C C T A T G G T C T G A T C A G A A A T A C T G C T A C A T C T T G C A C A G A C C T T T A A G A A	4361
D	b	4436	t g c c t a t g y g g t c t g a t c a g a a t a c t g t c t g c a t c t c t g c a a g a g c c t c t a a g a a	4495

```

OY 4362 ACCACTGTACAAATATGACACTGACCTAGAGATTATTTGGGAAGGAATATATTGTTCCAGGCATC 4421
Db 4436 accactgttcaaatatgacactgacctagagattatttgggaaggaatattatgttccaggcatc 4555
OY 4422 CATCTTGGGAAGAGCAGCAGCTTTGAGCTCCATCAGAGGATTTTAAAGTATGTCATTT 4481
Db 4556 catcttgggaaggaagcagcagctttgagctccatcagaggatTTTAAAGTATGTCATTT 4615
OY 4482 GAGCAGGAAT 4491
Db 4616 gagcaggact 4625

RESULT 9
US-60-213-360-7991
; Sequence 7991, Application US/60213360
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Poly
; TITLE OF INVENTION: Identified Thereby
; FILE REFERENCE: GX-0014 P
; CURRENT APPLICATION NUMBER: US/60/213,360
; NUMBER OF FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 8347
; SOFTWARE: PERL Program
; SEQ ID NO 7991
; LENGTH: 7116
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO: 1094199.1
; NAME/KEY: unsure
; LOCATION: 3941-3993, 5899-5939, 6951
; OTHER INFORMATION: a, t, c, g, or other
US-60-213-360-7991

Query Match 97.7%; Score 4393.8; DB 53; Length 7116;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 4437; Conservative 0; Mismatches 54; Indels 19; Gaps 2;

OY 1 GCTGATGATGAGACTCAGAGACCAATAAATAAAGTCTGTGAACATCCTTTGACTGGTT 60
Db 116 gctgatagtgtgactcagagaccaataaataaagctgttgaaacatcctttgactggtt 175
OY 61 AGCCAGTTCGTCGATGTATATTTCAGATGAGTGGATTAGAGAAAACCTTGATCCACTGCC 120
Db 176 agccagtttcgtcgatgtatatttcagatgagtggtattagagaaaaccttgatccactgcc 235
OY 121 CAGTATTTCAGCAAAAACCAAAATTCGATGTGATCTCAGAGCAAGGCTTACTGTAG 180
Db 236 cagtatttcagcaaaaaccaaatttcgattgtgatactccagacaaggcttactgtag 295
OY 181 TGGTGA AAAACGAGACGAGCGGAGCAGAAAAGTAAATATATTGAAAGATTGGCTGAGCTGAT 240
Db 296 tggtgaaaaacgagacgagcgaggagcaagaaagtaatatattgaagatttggctgagctgat 355
OY 241 ATCTGCCAATCTTATGATATTGACAATTTTCATGTTCAAACAGATTAATGTGGCATTTT 300
Db 356 atctgccaatcttattgatatttgacaatttcagtccaaccagataaattgctgatttc 415
OY 301 AAAGCAAAACAGTAAACAGATTCGTCAATTA AAAAGCAAGCAAGAAAACCTATTTCCAATGA 360
Db 416 aaagcaaacagtaaacagatctcgtcaattaaaagcaagcaagaaaacctattccaatga 475
OY 361 TGATGATGTTCAAAAAGCGATGTATCTTTCACAGGCGAGGAGATTATTTGATTAAGACTC 420
Db 476 tgatgatgtttcaaaaagcgatgtatcttccacagagcgagagattatttgaataaagactc 535

```

OY	421	CTTGAAGCCGCTTTTACTTTCAGGCATTTGAGTGGTTTCTATTTTGTTGTAATCGAGAGCC	480
Db	536	cttagacgcgcttctaacttaacttaacgcatlgbgtttccatttllgfygtaatcgagcgcg	595
OY	481	AAACATTGTATTTGTATACAGAAAATGTCAACAAATACCTGGCAATTAAGCAAGAGACT	540
Db	596	aaacattgtaatlgtatacgaanaaaatgccaacaataccgtaataataagcaagagacct	655
OY	541	GGTTAAACACAAGTCTTTTACAAATATCTTTAATGAAAGACACAAAGATTTCTTTAAGAA	600
Db	656	ggttaacaacaagtggttaacaatacttaactaigaagaagcagaagaagtttcttaagaa	715
OY	601	TTTTCACAAATCTACACTTAATGAGACTTTCTGGACAAATGACCCCAAGACAAAAG	660
Db	716	tttaaccaaaatctacaglttaigtatcttcctgycnaaatgagaccaagaacaaaag	775
OY	661	CCATACATCTTTAATTTGGCGTATGTGATGAAAACACCACTGATTTCTTGGAAGACATATA	720
Db	776	ccaatacalttaattgcgcgtatglttcgtatgaaaacaacaatgtaatlccggaagacaata	835
OY	721	CGCCAGTCTCGAAAATGCGCCAGAGAAATGAAACAATGCACTGTTCTTTGCCCTGTCAACC	780
Db	836	cgcgcagtcctgnaaatgycgcagagataigaacaatgcaatgcttgcttgccctgtccagcc	895
OY	781	ACGAGCTATGATGAGGAGGAAGGGCAATTTGCCAATCTTGTATGATCTGTGTGGCACCGC	840
Db	896	acgagctatgatgtaggaagaagggaatcttgcaatcttgatagatcgtlgtfygacgcgcg	955
OY	841	CATTACTACAGGAGAAAGACATTTTCCATCAAAACCCCTGAGACTTTATTACAGACATGA	900
Db	956	cattactacagggaaagaacaatlccaaccaaaccttgagagcttataaccagaatga	1015
OY	901	TCCTTTCAGGAAAGTTGTCAATATATAGATCAAAATTCCTAGATCCTCATAGAGCCTCG	960
Db	1016	tcttccaggaaggtctgtaataatagataagatacnaatcaactgagatcctccaatgagcctcg	1075
OY	961	CTTTGAAGATATATACTCGAAGGTGTATTCAGATTTTACTCTAAATGATGATGGGCAGTC	1020
Db	1076	ctttgaagaatataactccgaagaagcgtatccaagatcttcttagtcttaaatgtaggcagctc	1135
OY	1021	ATGTGTCCAGAAAGTCCATCATCAACAAAGCTTATCTTAATGCGCCATGAGAAACCCAGT	1080
Db	1136	atggtcccgagaagaagcgtactataccaagaagcttatacttaatgscatlgcagaaacccagct	1195
OY	1081	ATATCGATTTCTCGTGGCTGTATGGAACCTATAGTGACTGACAGACAAAAGCAACTCTT	1140
Db	1196	atactgattctcgtctggtcgtatgaaactatgtagtgcacagacgaataaaagcaactctt	1255
OY	1141	CCGAAATCTCGTAAACAATATGATGACATGGCTTTGTCTCAACCCCACTTCCTTCAGAGAGA	1200
Db	1256	ccgaaatccctgtaacaaatgatactgcacatggtcttcttcaaccacttcccttcagagaga	1315
OY	1201	ACAGATGATATAGACCAAAACCCAAATCCTGTTGGACAAAGGATTTAGACCACTATGGC	1260
Db	1316	acagaatgataatagaccnaaacccaaatccgtltgtaacaaggtatlagaccactatgyc	1375
OY	1261	TGAGATGCAACAGTTCGGTATAGCGGGCATGAGTATGTCGCCAAACCAAGCTTACAGATGCC	1320
Db	1376	tgagatgcaaacglttcglttaggcgcatgtagtatlgtcgcnaaaccaaggtcttaagatgyc	1435
OY	1321	GAGACAGAGGGCTATGAGTGTGGCAACACCTAGACACACAGGGCAGATGATGGAGCTAG	1380
Db	1436	gagacagagggcctatggtcttggcagacccttaagcaaccaaggtgcagatgagatgagactag	1495
OY	1381	GTAATGGGGTTCACGTAAACATACCTTCATTTGACCCCTGGGGCCAGGCAATGCAATCACATC	1440
Db	1496	gtaatggggttccagtaacataagcttcatltgaccccttggccaagcatgcaatcacatc	1555
OY	1441	TTCTTACCAGAAACAACATATGGGTGCMAAATGATACCCGCCCAATGGAGATGCTGG	1500
Db	1556	ttcctaccagaacaacaactatggtgtccacatgtagtgcctccccaactgtaggtctcgtg	1615

Qy	1501	TCCTGCCCCAACGACGACGAATATCATGATTTCTCTGTAATCGTGGAGTCCAAAGAT	1560
Db	1616	tccttgcccaaaccaagcagaataatcatgtatcttcctctg taactg tggagtgccaagaat	1675
Qy	1561	AGCCTCACATAGTTTCTCTGTTGCGAGTGTGCACTCTCCCATGGCATCTTCTGGCA	1620
Db	1676	agccccaataagttcttcctctgtgcaggtgtgcactctccca tgcactcttcyga	1735
Qy	1621	TACTGGGAACACACAGCTTTTTCACGAGCTCTGCAGTGTGCCCTGCAGACATCAGTGAAG	1680
Db	1736	tactyggaaaccaagctcttccagcagctctccag tgcctgcagacccaatgtaag	1795
Qy	1681	TGTGGGACCTTCCCTTTTATCTACTCTGTTCATCCACGGCCCCAAATTTGATTACTCTCC	1740
Db	1796	tgttgggacctctcctttctatctactctgtcatccacagccccaaat tgaatactctcc	1855
Qy	1741	CAATATGAATTTTACCCACACCAAGTAAGTAAGCAATTCAGATTTCCAGAGTCCCTTGGG	1800
Db	1856	caat tgaata tta cccaaccaagtaagtaagcaatcagaat tccaagatctccctcyg	1915
Qy	1801	CTTTTATATGACACCAAAATCATGAGAGATTTCAATGTCTGCATTAATTAACAGAGATCA	1860
Db	1916	ctttattatggaacccaataatccag tggaggttcaat tgyt tca tgcataaatagcagatca	1975
Qy	1861	CCTCATGACCAAAAGAAATGAAGAGACAGTGTGTAAGGGGGGAGAGAAATCAAAGGGTCC	1920
Db	1976	cctcaag tgaacaagaagaagtaagagagcag tgyt tggaggggcagaagatccaagggtcc	2035
Qy	1921	TTTGGAAAGCAAGGTGATATAAAATTTACTGCGATTACTTACCCTGTTCTTGATGACCG	1980
Db	2036	tttggaaagcaaaagtgataaaaaaatctac tgcagttacttacctgtctctctgtatgacg	2095
Qy	1981	GGGTATATCCCTCTTGACCAACTCCCCCTAGATTCATCTTAAAGATCTTGCTTTAG	2040
Db	2096	gggtatctccctctctgtgcaaatcccccctagat tcaag tgttaagaatctctgttag	2155
Qy	2041	TGTCAACGACCCCTCTGGAGTCTCTCTCTTACATCTGAGAGAGATCCCTACATCCAA	2100
Db	2156	tgtcaacagccccctctgagtgctctctctctacatct tgaaggagatcctcttaactcaa	2215
Qy	2101	TATGATAGGGTCACTGTTACAAAGAAAGACCCGGATTTTTCACAAGTTGCTGCGAATAGG	2160
Db	2216	tatgata tggactctgttaacaagaagacccggaat tcttgcacaag tctgcgcagaatcyg	2275
Qy	2161	GAATTCAACACACCTAGGTAGCAGCAAGATTACTGCAACAAGCCATGGGAAAGACACAGAG	2220
Db	2276	gaattcaaccagct tgaaggttagccaagat tta ctgcagaagccactctggaaaagacacagcag	2335
Qy	2221	TATAACTCTTGTGGGGACGGAAATGTTGTCAAGACGAGACGCTAAGTCTTAAGACAA	2280
Db	2336	tataactctctgt tgggagcgaataatgt tgcacagcagcagcagtaag tcttaagaagaa	2395
Qy	2281	GGAGATAATATGCACTTCTTGTAGATATCCCTGCTGACACAGGATGATCCTAGTATGCACCTTC	2340
Db	2396	ggagataata tgcactctctatgatactctgc tgcagaag tga tctctag tga tgcactctc	2455
Qy	2341	TAAACAATCAGGCCCAAGGGAAGAGAGGACATTAATAATGAGTCAAGTGCACAGCTC	2400
Db	2456	taaaagactcaagcccccaag tgaaggaag tgcgaataa aa tgcgtcag tgcacagctc	2515
Qy	2401	CACCACTTCTAGCTCAAGTCAAGAGAAAGACCTTAATAATTAAGACAGACAGATGAAGA	2460
Db	2516	caacctctctagctcaag tcaaggaagaccccttaaat ttaagacagagacaag tgaaga	2575
Qy	2461	GGGATCTGGAACATTTGCATATCTAGATGCTATTTCTTGGTGATCTGACTAGTTCTGACTT	2520
Db	2576	gggactctggaact tgaataatctagaatctatctct tgc tgcactag tttc tgcact	2635
Qy	2521	TTACAAATAATTCATATTCCTCAAAATGGTACTCATCTGGGAGCATTAAGCAACAAGGTTTCA	2580
Db	2636	ttacaataatccatactccccaat tggtagtca tcc tggggactaagcaacag tgtttca	2695
Qy	2581	AGGAATCAATTCCTGTGGGTTTGAAGAATTCACAGTCTTGACAGTCTATTTGCTCTCCATA	2640

Db	2696	aggaactaatctctctggttctgtaaaagtctacacagctctgctgaactctatctgctccata	27555
Qy	2641	TAACCGAGCACTGCTCTGGATAGACCTGTTTCTGTGGCTCAAGTCTTCCAGTAAAAA	27000
Db	2756	taaccgagcaagtctctctgatacagccctgttctcttgctcaagctccctccagtaaaaaa	2815
Qy	2701	TATCAGTGTCTTCCCAATGTTTACCAAAAGCAACCATGTTGGGGGGAATCCAAAGATAT	27600
Db	2816	tatcaagtctcttcccacatgcttaacaaagcaaccacagtttgggttggaaatcccaagaatgat	2875
Qy	2761	GGATAGTCAGGAAAAATTATGGCTCAAGTATGGGTGGGCCAAACCGAATGATGCTGTGAC	28200
Db	2876	ggaatagtaagaanaaatatagctcaagtaatggttgggtccaaaccgaaatgtaactgtgac	29355
Qy	2821	TCAGACTGCTCTTCGACAGACAGCTGGGGCTTACCAACTCAAGGCGCGAGAAATGGAAAC	28800
Db	2936	tcagactctcttcccagaaagctgggtcttaacaaactcaaaagccggagaaatgaaac	29950
Qy	2881	TATGATTTCAAACTCCATGGGAAGACCGAGGAGAGATTTTAACTACTTTTACCCAGAC	29400
Db	2996	tatgaatctcaaacctccatlggaagaccaggaaggtatataactcttcaaccagacc	30555
Qy	2941	TGCACGTGGGTGGCTCTATTCCACATTGCTTCTTGCGTCTAATAGCATACAGGTGCGAG	30000
Db	3056	tgcacttgggtgggtctctatctccacatgtcctctcttggtctaaatagcatctcaagtgtag	31155
Qy	3001	ACCAATATTGCAACAGACAGACAGATGCTTCAAAATAGAGCGCTGGGGAATCCCATAGG	30600
Db	3116	accagtatctgaaagaacagcagcagatgctctcaaatgaaagcctgtgtgaaatctcccatgg	31755
Qy	3061	AATGGGGCTAATCCCTATATGGCCAAAGCAGCATCTTAACCAACTGGGTTCTTGCCCCGA	31200
Db	3176	aatggggctaatctccctcatlgccaagcagcagcatctaaaccacatgggtgtctctggccga	32355
Qy	3121	TGGCATGTTGTCATGGAAACAAGTTCTCATGGCATCAAAATAGGCGCTCTTCTTAGAA	31800
Db	3236	tggcatgtgtctccatggaaacaagtcttccatggcactcaaaataggtccctctcttagaa	32950
Qy	3181	TTCCCTGGATATCTTTTGGGGCCACTTTCACACTGTGAAGGCCAGAGTGAAGAAAGGC	32400
Db	3296	tcccttgatgatactgtgttgggccactcttccaaacttgaaagccaggaagtgaagaaagtc	33555
Qy	3241	ATTATTTGGACCAAGCTGCACACTCTTCTCAGAACACAGATGGACAGGCTTGGAAAGAAAT	33000
Db	3356	attatlgacaagcttgacaactcttcttcagcaacaagaatgccaaggtcttgaagaat	3415
Qy	3301	TGACAGAGCTTTGGCAATTCCTGTAAGCTTGTCATTCAGGACAGAGCAATTAGACCCCAACA	33600
Db	3416	tgaagagcttgggcaatttcctgaactctgcaactgtgcaatcgaaggaagcttatagaccaca	3475
Qy	3361	GGATGCTTTTCCAAAGGCCAAGACAGACAGTAAATGATGATCAGAAAGCAGATTATATGG	34200
Db	3476	ggatgcttcttccaaaggccaagaagcagataatgatactgaaagccaggaatctatctg	35355
Qy	3421	ACAGACATACCACACAGAGGGGCGTCCAAATGCAAGAGAGCTTTCATCTTACAGGCATTC	34800
Db	3536	acagacataccacagcaaaaggtcctccaatgcaagaagagcttcaactcttcaaggcaatc	35955
Qy	3481	ACCATCTTTTAAGTCTATGATGATCAGATGAACACGAGAGCAATTTTCTCTTCAAGG	35400
Db	3596	accatcttctaactctatgatacagatgaacccagcaagcaatttctctctcaagg	36555
Qy	3541	AATGACACCCACAGACCAAACTCATGAGAACCCCGCAAAACACCCCAAGCAACTTAGAAT	36000
Db	3656	aatgacacccagagccaactcatctgaaacccggaaacacccccaagcaacttagaat	37155
Qy	3601	GCAGCTTCAGCAGAGGCTTCAGAGGCCACAGAGTTTGTGATTCAGACCCGACAGCACTTGA	36600
Db	3716	gcagcttcagagaggtctgcaagggtccagcaagtttctgaaatcagagcgaacaggtctga	37755
Qy	3661	ATTGAAAATGGAAAACCTTACTGCTGTGGTGTGCTGCGGTGATGAGCCTTATGATGACAC-	37199

Db	3776	attgaaatctgaataaaccttactgtcgttggtggctgcgtgcatgagccatacgaccc	3835
Qy	3720	-----CCACAGACGGGTTTCTTAATGCTCAAAATGGTCGCCAACGACGACGAGA	3768
Db	3836	ccaaggtagctcccagaaggtttctcttaagtctcaaatggttcgccacaacgacagaga	3895
Qy	3769	GCTGTGAATGCATCATCTTCGACAAACAGAGGGTGCGTATGATGATGACGACACACACA	3828
Db	3896	gctgtgaagtcatactactctccgaacacacagaggtgggtctatgatacnnnnnnnnnn	3955
Qy	3829	GCACGACGACGACGACGACGACGACGACGACG-----ACACGACGACGACGACG	3881
Db	3956	nnnaacagacagcacagc	4015
Qy	3882	AGCAAACCACGCGCTTCAGGCCAOCCTCCTAATGACTGCTTCCCAGCATGATGGCG	3941
Db	4016	aagcaaacccagagcttcacagccacctcttaaagtgcagctctccccacagatggc	4075
Qy	3942	TTTTGGCAGGACCCACATGCGCACAAAGCTTCCTCCGACAAAGTTTTCCATATCAACCAATT	4001
Db	4076	ttttggcagggccccacaatlgccacaagctctctccgcgaacagttccatatcaaccaatt	4135
Qy	4002	ATGGAATGGGACAAACAAACGAGTCCAGCCTTTGGTGAATGTCCTAGCTCTCCAATGCA	4061
Db	4136	atggaaatgggacaaaacacacagatccagctcttggtagaagtctagctctcccaaigtcaa	4195
Qy	4062	TGATGTCCTCAAGAAATGGTGCCTCCAGAAATCCATGATGCAACACCGCGAGGCTGCAT	4121
Db	4166	tgatgtctgcaagaatgggtgcctctccacagaaatcccatgtagcaacaccgcagagctgat	4255
Qy	4122	CCATCTATCAGTCTCTCAGAAAATGAAGGCTGGCCATCAGAAAAATTTGGCCAGAACAGCT	4181
Db	4256	ccatctatcagctctccagaaaaatgaaggtctgcccacacgaatcttggccagaaacagct	4315
Qy	4182	CCTTTCCACACGACGATTTGGCCACAGGGGAATCCTGCAGTATGATGATGGGACGA	4241
Db	4316	cctttcccaacagcagtttggccacacagggaaatccctgcagtgataagtaagtagcgca	4375
Qy	4242	TGAATGGCAGCAGTGGTCACATGGGAGAGATGAAATGAAATGAAACCCCATGCTTGCGCA	4301
Db	4376	tgaatggcagcagtggtccatagtggacagatggaaatgaaacccatgcccattgctggca	4435
Qy	4302	TGCCATATGGGTCCTCGATCAGAAATACTGCTGACATCTTTCGACACGAGACCTTTAAGGA	4361
Db	4436	tggcataatgggtccctgcatacgaataatctgcycacacagagacctctaaggaa	4495
Qy	4362	ACCACTGACAAATGACACTGACACTAGGATTTATTTGGGAAGAAATATTGTTCCAGCATC	4421
Db	4496	accaactgacaaatgaaacatgcacatgatatcttggaaagaaatcatctgtctccagcatc	4555
Qy	4422	CATCTTGGACAGAAAGACACGACTTTGAGCTTCATCAAGGGTATTTTAAAGTAGTGCATTT	4481
Db	4556	catcttggaaagaaagacacagctttagagctccatcaagagtatctttaagtatgcatctt	4615
Qy	4482	GAGCAGGAAT 4491	
Db	4616	gagcaggaact 4625	
RESULT 10			
US-09-440-612-1			
; Sequence 1, Application US/09440612			
; GENERAL INFORMATION:			
; APPLICANT: C. Frank Bennett			
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION			
; FILE REFERENCE: RFS-0042			
; CURRENT APPLICATION NUMBER: US/09/440,612			
; CURRENT FILING DATE: 1999-11-15			
; NUMBER OF SEQ ID NOS: 47			
; SEQ ID NO 1			
; LENGTH: 6754			
; TYPE: DNA			

```

RESULT 10 RESULT
US-09-440-612-1
; Sequence 1, Application US/09440612
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
; FILE REFERENCE: RTS-0042
; CURRENT APPLICATION NUMBER: US/09/440,612
; CURRENT FILING DATE: 1999-11-15
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
;
; LENGTH: 6754
; TYPE: DNA
;
;

```

```

; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: (184)..(4422)
;
US-09-440-612-1

```

Query Match	96.18;	Score 4322.2;	DB 18;	Length 6754;
Best Local Similarity	98.08;	Pred. No. 0;		
Matches 4438; Conservative	0;	Mismatches	8;	Indels 81; Gaps 3

QY	1	GTGTGATGAGGACTGCTCAGAGCCCAATTAATAATTAACGTGTGAACATCCCTTTGACGGTT	60
Db	99	gctggaaggaggactccagagaccataataaacctgcttgaacaccccttgactggtt	158
QY	61	AGCCAGTTGGCTGATGATATATTCAGATGAGTGGATTAGAGAAAACCTTGGATCCACTGGC	120
Db	159	agccagcttgctgactgtatatcttaacagatgagtgatagagaaaacttgaatccactgag	218
QY	121	CAGTGATTTACGAAAACGCCAATTTGCCATGTGATCTCCAGGACAAAGTTTACCTGCAC	180
Db	219	cagtgatcttcagaaaaacgcaaatgcccactgtaactccagacaagctcttaccctgcag	278
QY	181	TGGTGA AAAACGCGAGCGGAGCGAGCGAAGAAATATATTAATGAAGAAATTTGGCTCAGCGAT	240
Db	279	tggtagaaaaacgagagacgggagccaggaagaataatatatgaagaatctgctgactgct	338
QY	241	ATCTGCCAATCTTAGTGATATTTGACAATTTCAATGTCAAAACGATPAATGTGCGATTTT	300
Db	339	atctgcgaatcttagtgatatctgcaaatcttcaaatgccaacccagataaagtctgacttt	398
QY	301	AAAGGAACAGTAGACAGATAGCTCAATTAATAAGACACAGGAAAAACCTTTTCCAAATGA	360
Db	399	aaaggaacacgataagacagataacgtccaaataaaagacaggaagaaaaactatcccaatga	458
QY	361	TGATGATGTTCCAAAAGCCCGATGTATCTTCTACAGGGCAGGAGTTATTTGATAAACATC	420
Db	459	tgaatgatttccaaaaagccgactgtaactcttccagggccagggtatctgaataagact	518
QY	421	CTTAGGACCGCTTTACTTTCAGCGATTTGGATGTTTCTTATTTTGGTGAAATGACAGGC	480
Db	519	cttaggacgcttcttacttccaggaattgtagtcttcttcttcttctgtagtaatgcagaagg	578
QY	481	AAACATGTATTTTATACAGAAAATGTCACACAATFACTCCATATTAAGCAAGAGACCT	540
Db	579	aaacattgtattctgtatccagaaaaatgccaacaatacttcgaataaagaagaagact	638
QY	541	GTTTAACACAAAGTGTTCACATATCTTACATGAAGAAGACGAAGATTTTCTTAAGAA	600
Db	639	ggttaacacaagtgcttcaaatcttcaatcttaagaagaagagaagatttcttaagaa	698
QY	601	TTTACCAAAATCTACACTTAATGAGAGTTTCTGTGACAAATGAGCCCCCAAGCAAAAAAG	660
Db	699	tttaccaaaactctacagcttaattgagtttccctgtgcacaatggaaccocaaagacaaaaag	758
QY	661	CCATACATTTTAATGCGCGTATGTGATGAAGAAACCCACATGATATTTGTGAACACATTAAT	720
Db	759	ccatacatcttaattcgcggtatgttgaagaaaacccacatgatacttcgtaagaacataaa	818
QY	721	CGCCAGTCTCGAAATGCGCCACAGATATGAACAACATGCACTGCTTGCCTCTGTCCAGCC	780
Db	819	cgcagatctcgaatgcgcgcagagatatgaaaacaatgcagtgcttgcctgtctcagcc	878
QY	781	ACGAGCATATGATGAGAGAGGCGGAAGATTGTGCATCTTGTATGATCTGTGTGGCAGCCG	840
Db	879	acgagcatatgtctggaaggaagggaagattcgcaactcttgatgatactgtgtgacacgg	938
QY	841	CATTACTACAGAGAAAGAACATTTCCATCAAAACCTGAGAGAGTTTATTAACACGACATGA	900
Db	939	cattactaagaggaagaagaacattccaataaacctcgagagctttataatccagacatga	998
QY	901	TCCTTTACGGAAGGTTGTCAATATAGATACAAATTAACGTAGATCCTCCATGAGCCCTGG	960

Db	999	tccttcagggaaaggtgtgtcctaataatagatacaaatctcaatctgagatccctccatgagggccttg	1058
Qy	961	CTTTGAAATATATATCCGAAGGTGTATTACAGAGATTTTTTTTATGCTTAATGATGGCAGTC	1020
Db	1059	ctttgaagataataatccgaaggtgtgattccagagatcttttaagctaaatgatgaggagct	1118
Qy	1021	ATGGTCCAGAAACGTCACTATCAAGA-----AGC	1050
Db	1119	atggtcccaagaagctcactactcaagaagttaccagtgatgysatattctcccaacagc	1178
Qy	1051	TTATCTTAATGGCCATGTGAGAAACCCCGTATATCGATATTCGTGGCTCATGTGAACATAT	1110
Db	1179	ttaacttaatgagccatcgagaaaccccaaglatatcgaatcttcgttcggtgatggaactat	1238
Qy	1111	AGTGACTCCACAGACAAAAAACCAACTTCTCCGAATCTCTTAACAATGATGCATGTG	1170
Db	1239	agtgcctgcaagagacaaaaagcaactcttcgcgaatctcgttaacaaatgatgcacttgg	1298
Qy	1171	CTTTGTCTCAACCCACTTCTTTCAGAGAGAACAGATGGATATAGACCAAAACCCAAATCC	1230
Db	1299	ctttgtccaaacccctctctccagagagaaacagatgatalatagaccaaacccaatcc	1358
Qy	1231	TGTTGGACAAGGGATATTAACCAACCATGTGGTGGATGGATGCAACAGTTCGGTTAGCGGCATGAG	1290
Db	1359	tgcttgaaagaaggatctagacaacctaagctgcatgcaacagttcgttgaagcgagag	1418
Qy	1291	TATGTGCGCAAAACCAAGGCTTTACAGATCCCAAGACAGCGGCGCTATGTGGTCGACAGCC	1350
Db	1419	tatgtgcgaacaaacaaagcttaacagatgtcgagacagcaagggcctatggtcttggagagccc	1478
Qy	1351	TAGCACCAACAGGGCAGATGATGAGCTTAGGTATGGGGGTTCCAGTAAACATAGCTTCAATT	1410
Db	1479	tagccaccaagygacagatgagtgagagctaggtatgaggggttccagtaacatagctcatc	1538
Qy	1411	GACCCCTGGGGCAGGCAATGACATACCACTCTTCTACAGAACAAACAATAGGGCGTCAA	1470
Db	1539	gacccctgggcgaagcaatgaacaaacaaactcttccaaagaaacaaactatagcttgcacaa	1598
Qy	1471	CATGATGATGACCCCCACATGGGAGTCCAGTGTCTTGCCCAACCAACGAGATATCATGAT	1530
Db	1599	catgagtagagcccccaacttgagagtcctggtcttgcgcccccaacgaagaaatatacatgat	1658
Qy	1531	TTTCTCTGTGTAATCGTGGGAGTCCAAAGATAGCCTCACATCAGTTTTCTCTGTGGCAGG	1590
Db	1659	tctctcctgtaactcgttgagagttccaaagatatagcctcaatcagttctctcgttgagag	1718
Qy	1591	TGTGACATCTCCATGAGGATCTTCTGGCAATATCTGGGAACACAGCTTTTCCACACACTTC	1650
Db	1719	tgtagaacttcccaatgagatctctcgtgaataactctggagaaacacagacttctccagcaactc	1778
Qy	1651	TCTCAGTCCCGTCGAAGCATCATGAGAGTGTGGGGACTCCCTTTTATCTACTCTGTC	1710
Db	1779	tctcagtgacctgcaagccaatcagtgaaagtggtggagaaacttcccttatactactcgttc	1838
Qy	1711	ATCACACAGGCCCAAAATTGATTAACCTCTCCCAATATGATATTTATCCCAACCAAGTAAAGT	1770
Db	1839	atcaccaagcccccaaatgtgataactctcccaatatgataataatcccaacaaagtaagtc	1898
Qy	1771	AAGCATACAGATTTCCAAAGATCTCTGTGGCTTTTATTGGACCAAAATCCAGTGGAGAG	1830
Db	1899	aagcaatcaagatgattccaaagatctcctctggtcttattatgagcaacaaatccagtgagagag	1958
Qy	1831	TTTCATATGTACGTAGTAATATACAGAGATCACCTCAGTGTACAAACAAAGTAAAGAGACAG	1890
Db	1959	tccaatgtgtaagtcataaataagcagagatcacctccagtgacaagaanaagtaagagagagcag	2018
Qy	1891	TGTTAGGGGGGAGAGAAATCAAGGGGTCCTTTTGGAAAGCAAAAGCTATATAAAATTAATCT	1950
Db	2019	tggttagggggagagaaatcaaaaggggtcctttgaaagcaaggtcataaaaaaatattact	2078
Qy	1951	GCAGTTACTTACCTGTTTCTTGATGATCCGGGGTCAATTCCTCTTGAACCAATCCCCCCCT	2010
Db	2079	gaagttaactacacttctcttcgtatgaaacgggggtcaatccctcttgacaaactccccct	2138

Db	4254	tttggcaggaacagctcccttlttcccgacgagtttggccaccaggggatacttcgact	4311
OY	4225	GTATAGTATGGTGCACATGAAATGGCAGCAGTGTTCACATGGAGCAGATGAATGAACCC	4288
Db	4314	gtatagtatgtgtgacatgaatgagtcagcaggtgtgtcacaatgagacagtgaaatgaaccc	4377
OY	4285	CATGCCCATGTCCTGGCATTCGCTATGGGTCCTCGATCACAATAATTACTGCTGACATCTCTGCAC	4344
Db	4374	catgtccaatgtctcgtgcattgtcctatgggttccgtgacagaataactgtctgcaatctctgcac	4433
OY	4345	CAGGACCTCTTAAGGAAACCACTGTACAAATGACATGCACTAGAGATTATTGGGAAGAA	4400
Db	4434	caggacctcttaaggaataacccctgttacaatgaacatcgtcactatgtattatltgggaaggaa	4499
OY	4405	TCATTTGTCACGSCATTCATCTTGGAAAGAAAGACACAGCTTTGAGCTCCATCAAGGGTAT	4466
Db	4494	tcatgtttctcagcagcatccatctltggaagaagacagccttggatgcctcatcaaggttat	4555
OY	4465	TTTAAGTGATGTCATTTTGAGAGGAAT	4491
Db	4554	tttaagtgatgtcatctltggaaggaact	4580

```

RESULT 11
US-09-440-612-3
: Sequence 3, Application US/09440612A
: GENERAL INFORMATION:
: APPLICANT: C. Frank Bennett
: APPLICANT: Lex M. Cowsett
: TITLE OF INVENTION: ANTISENSE MODULATION OF SRC-3 EXPRESSION
: FILE REFERENCE: RTS-0042
: CURRENT APPLICATION NUMBER: US/09/440,612A
: CURRENT FILING DATE: 1999-11-15
: NUMBER OF SEQ ID NOS: 49
: SEQ ID NO 3
: LENGTH: 6754
: TYPE: DNA
: ORGANISM: Homo sapiens
: PEATUNE:
: NAME/KEY: CDS
: LOCATION: (184)..(4422)
: US-09-440-612-3

```

Query Match	96.18;	Score 4322.2;	DB 18;	Length 6754;
Best Local Similarity	98.08;	Pred. No. 0;		
Matches 4438; Conservative	0;	Mismatches 8;	Indels 81;	Gaps 3;

[illegible]

OY	361	TGATGATGTTCCAAAACCCGATCTGATCTTCTACAGGCGGAGGAGTTATGATTAAGAGCTC	420
Db	459	tgaatgaatttcaaaaaagccgaatgatctcttcaaaagggcaggagattatgtaaaagactc	518
OY	421	CTTAGAGACCCGTTTACTTTCAGGCACTTGGATGGTTTCCATTTTGATGGTAATCGAGAGCC	480
Db	519	cttagagaccgcttctaacttaaggcatcttgatcttccattcttgatggaatcgagaacg	578
OY	481	AAACATTTGATTTTGTATCAGAAAAATGTCACACATATACCTGCATATTAAGCAAGAGACCT	540
Db	579	aaacattgatcttctgtatccagaaaatgtcaacaatacctgcataataagcaagagacct	638
OY	541	GCTTAACACACAGTGTTTACATATCTTTACATGGAAGACACAAAGATTTTCTTAAGAA	600
Db	639	gcttaacaacaagtgatttacaataatcttacaatgaagaacagaaagatttctttaaaga	698
OY	601	TTTACCAAAATCTACACTTAAATGAGTTTCCTGACAAATGAGCCCAAGACAAAAGAG	660
Db	699	tttaccaaaatcttcagtttaatgagtttcccttgagacaatgagaccacaagaacaaaaag	758
OY	661	CCATACATTTTAATTTGCCGTATGTTGATGAAAAACCCACATGATATTCTGGAAGACATAAA	720
Db	759	ccaatacttaattgcccgtatgttgatgaaaaacacacatactgtatcttcygaagacataa	818
OY	721	CGCCAGTCTCTGAATTCGCCGACAGATATGAAACATATGCAGTCTTTGCCCTGTCTCAGCC	780
Db	819	cgcacagctctgaaatctgcgccaagagataatgaacaatgacagttcttgccttgcctcagcc	878
OY	781	ACGAGCATATGATGAGAGACGGGGAAGTTTGCATTTTATGATCTGTGNGGCGACGGC	840
Db	879	aagagcatatgctggaaggaaggggaagatcttgcaatcttctgtaagatctgtgtgacgcg	938
OY	841	CATTACTACAGAGAAAAGAACATTTTCCATCAAAACCCTGAGACCTTTATACAGACATAGA	900
Db	939	cattactacagaggaagaagaacatttccatccaacccctgagagccttattatccagaacatga	998
OY	901	TCTTTACAGAAAAGTTGTCAATATATGATACAAATTCATCAGATCCTCCATGAGCCCTGG	960
Db	999	tctttccagaaaggtctgtaataatagataaacttcaatcagatcctccaatgaagccctg	1058
OY	961	CTTTGAGATATTAATCCGAGAGTGATATCGAATTTTATCTGTAATATGATGGCAGATC	1020
Db	1059	cttgaagaataataactccgaagtgctatccagagatttltttagtcaaatgatatgycagctc	1118
OY	1021	ATGTCGCCAGAAAAGTCATCATCAAGA-----AGC	1050
Db	1119	atggttccagaagaagcttcaactataccaagaagtctccagtgatggtatatttcccacaacgc	1178
OY	1051	TTATCTTAATGGCCATCGAGAAAACCCACGATATATCGATTTCTGTTGGCTGATGAACTAT	1110
Db	1179	ttatcttaattatgccaatgcagaaaaccccgatatatcgattcttcgltgcgtatgagactat	1238
OY	1111	AGTGACATGACACAGACAAAAGCAAACTCTTCCGAATTCCTGTAAACAAATGATCCGACATGG	1170
Db	1239	agtgactatgcaagaacaaaagcaaaccttccgaaatctctgtaacaatatgacgcacatg	1298
OY	1171	CTTTGTCTCAACCCCACTCTCTTCAGAGAGAACGAATGATATAGACCAAAACCCAAATCC	1230
Db	1299	ctttgtcttcaaccacacttcccttcaagagagaacaagaatgatatagaccaaaccacaatcc	1358
OY	1231	TGTTGGACAAAGGATATGACCACCTATAGCTGGATGCAACAGTTGCGTAGGGCGCATAG	1290
Db	1359	tgtttgacaagagattagaccacactatgctgtgatgcaacagttcgttagagcggatgag	1418
OY	1291	TATGTGCGCAAAACCAAGGCTTAACAGATGCCGAGACAGCAGGGCTTATGGCTTGCAGACCC	1350
Db	1419	tatgtgcgcaaaacaaaggtcttaacagatgcccgaagcagcaagggcctatggtcttgcagaccc	1478
OY	1351	TAGACACACAGGCGCATGTGATGAGTACATGATGTGGGGTTCTCAGTAACATAGCTTCATT	1410
Db	1479	tagcaaccaagggcagatgtgtgagcttagatgttgggggttccagttaacataagcttcaat	1538
OY	1411	GACCCCTGGGGCAGGCAATGCATACATCTTCTTAACGAAACAACATATGGGCTCA	1470

[illegible]

[illegible]

```

: PRIOR FILING DATE: 1999-08-30
: NUMBER OF SEQ ID NOS: 9796
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8947
: LENGTH: 7888
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-652-123-8947

Query Match      95.9%; Score 4313; DB 25; Length 7888;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 4436; Conservative 0; Mismatches 10; Indels 87; Gaps 3;

OY 1 GCTGATGCTGAGCTCAGAGACCAATAAATAACTGCTTGACATCCCTTTGACTGTT 60
Db gctgatgctgactcagagaccataataaactgcttgaacatcccttgcctgctgct 156

OY 61 AGCCAGTTCGTGATGATATATTCACATGATGCTGATTTAGAGAAACTTGATGCCACTGGC 120
Db agccagctgtcgtatgatatccaatgaatgagcttgaagagaactggaaccactgacc 216

OY 121 CAGTATTCACAGAAACCAAAATTCCTCATGTCATCTCAGAGACAAGTCTTACCTGCAG 180
Db cagtattccacgaaaccaaattcccatgtcctcagagacaagtcttaccctgcag 276

OY 217 cagtgatccagaaaaacgcaaatgtccatgtagatctccagagaaggtcttaccctgcag 276
Db 181 TGGTGAAGAAAGCGAGCGGAGCAGGAAAGTAAATATATTTGAAGAAATTCGCTGACTGAT 240
Db 277 tggtagaaaacggagcgggagcaggagaagaataatataatggaagaattgcctgactgat 336

OY 241 ATCTGCCAATCTTACGTATTTGACAAATTTCAATGTCAACACAGTAAATGTGCATTTT 300
Db 337 atctgccaatcttgatgatattgacaattttcaatgtcaaacagttaaattgcgacttt 396

OY 301 AAAGGAAACAGTAACAGATACGTACGTAATTAAGAGAGAGGAAACATTTTCCATGTA 360
Db 397 aaaggaaacagtaagacagatacgcgaataaagaagaagaataaacatcttccaatga 456

OY 361 TGATGATGTTCCAAAAGCGGATGTATCTCTTCAACAGGCGAGGAGTATTTGATTAAGACTC 420
Db 457 tgatgatgttccaaaagcgatgtatctcttcaaacaggcgaggtatgttgaataagactc 516

OY 421 CTTAGACCGCTTTTACTTTCAGGCAATTTGGAGTCTTCCATTTTGTGGTGAATCGAGAGGC 480
Db 517 cttagaccgccttcttaccgcagcatgtagtgccttccatttgcgtgaaatcgagacgg 576

OY 481 AAACGTTGATTTTGATGAGAAATGTGCAACAAATACCGCAATTAAGCAAGAGACCT 540
Db 577 aaacgttgattttgatgagaaatgtgcaaacaaataccgcaataaagcaagagacct 636

OY 541 GGTATTAACAAGTGTTTACATATCTTACATGAAGACAGCAAAAGCATTTTCTTAAGNA 600
Db 637 ggtatataacaagtgtttacaatatcttacaatagaagaagacgaagaagatttcttaagaa 696

OY 601 TTTACCAAAATCTACAGTTAATGAGATTTCCTGCAAAATGAGCCCCAAGAACAATAAAG 660
Db 697 tttaaccaaaatctacagttaatgagatttccctgcaaaaatgagccccaaagaataaag 756

OY 661 CCATACATTTTAATTCGCCGATGTGTATGAAAAACACCACTGATATTTCTGGAAGACATAA 720
Db 757 ccatacatcttaatttcglatgtgtatgaaaaacacacaatgatactctlgyaagacataaa 816

OY 721 GCCCAGTCCCTGAATAGCGCGCAGAGATATGAAGAACATGACGTCTTGCCCTGTCAGGC 780
Db 817 gcccagtcctgaaatagcgcgcagagatagaaacaatgcagtgcttgccttccagcc 876

OY 781 ACGAGCTATGATGAGAGAAAGGGAAGATTTGCCAATCTTGTATGATCTGTGTGACAGCGCG 840
Db 877 acgagctatgatgaggaaggggaagatttgaacatctgtatgactcgtgtgcaagcgcg 936

OY 841 CATTTCTACAGAGAAAGAAACATTTTCCATCAAAACCTTGAGAGCTTTATTTACGACATGA 900
Db 937 catctctacaggaagaagaacatcttccatacaaccctgagagcttataccaagacatga 996

```

```
QY 901 TCTTTACAGAAAGCTTGTGATATATAGATACAAATTTCAGTAATCTCCATGAGGCTCG 960
|||||
Db 997 tcttcaaggaaaggtctgtcaataataaatactacagatccctccatgagccttg 1056
QY 961 CTTTGAAGATTATTAATCCGAAAGGTGTATTCAGAGATTTTATGCTTAATGATGAGGACATC 1020
|||||
Db 1057 ctttgaagatacaatccgaagtgatctcagaatcttctagtctcaatgatagagcagtc 1116
QY 1021 ATGTGTCAGAAAGCTGACTATCAGA-----AGC 1050
|||||
Db 1117 atggtccagaagcgtcactatcaagaagtaccagtgatgggatatcttccccaacagc 1176
QY 1051 TTATCTTATGCGCATGCGAAACCACATATATGATTTCTGTTGGCTGATGGAATCT 1110
|||||
Db 1177 ttatcttaatgagcatgacgaatacccaagtataatcgtctcgtctgacgtatgaaactat 1236
QY 1111 AGTGAAGTACAGAGCAAAAGCAAACTCTTCGAAATCTGTAAACAAATGATGACATG 1170
|||||
Db 1237 agtgaatgacagacaaaagaacacacccgtccgaatccctgtaacaaatgatacgaca 1296
QY 1171 CTTTGTCTCAACCCACTTCTTCAGAGAGACAGATGATATAGACAAACCAATCC 1230
|||||
Db 1297 ctttgtcaaacaccactctccacagagatacagaatgatatagacaaacccaatcc 1356
QY 1231 TGTGAGACAAGGATTTAGACACCTATGGCTGATGCAACAGTTGCGTATGCGGACATGAG 1290
|||||
Db 1357 tgttgacaagaagatagaacacactatgagctgatacgacaagctcgtatgagcgac 1416
QY 1291 TATGTGCGCAAAACCAAGGCTTACAGATGCGGACAGACAGGCTTATGCTTGGCACAACC 1350
Db 1417 tatgtgcacaaacaaagccttcaagatgcagagcagcagcccttagtctggcagacc 1476
QY 1351 TAGACACACAGGAGAGATGATGAGACTAGTATGGGGTTCCAGTAACTAGCTCAT 1410
|||||
Db 1477 tagacacacagagcagatagatgtagctagatagtggtctccagtaacatagtctcat 1536
QY 1411 GACCCCTGGGCGAGGATGCAATACACATCTTCTACCAAGAAACACACTATGGGCTCAA 1470
|||||
Db 1537 gacccctggcagagcatgcaatccacactctccacgaacaaacataagctcaa 1596
QY 1471 CATGAGTAGCCCGCCACATGGGAGTCTGTGCTTGCCCAACACAGCAATATCATGAT 1530
Db 1597 catgtagtagcccccacacatggagatccgtgtcttgccccaaacagcagaatactgata 1656
QY 1531 TTCTCTCTTAATGCTGGGAGTCCAAAGATAGCTCATCATGATTTTCTCTGTTGACAG 1590
|||||
Db 1657 tctctctgttaatcgttggagatccaaagatagcccaatcagttctccctgttgcagg 1716
QY 1591 TGTGCACTCTCCATGGCATCTTCTGGCAATACTGGGAACACAGCTTTCAGACAGCTC 1650
Db 1717 tgtgaactctcccaatgacatctcttgcagcaatactggaacacacagcttctcagcagctc 1776
QY 1651 TGTCAAGTCCCTGCAAGCCATCAGTGAAGTGTGGGACTTCCCTTTTATCTACTGTGTC 1710
Db 1777 tctcagtgccctgcgaagccatcagatgaaggctggggacttcccttcttactactgtgc 1836
QY 1711 ATCACCAGGCCCAATTGGATTAATCTCCCAATATGATATTATACCAACCAAGTAAAT 1770
|||||
Db 1837 atcaccaaggccccaataatgtgataactctcccaataatgataatcccaaccaaagtaagtc 1896
QY 1771 AAGCAATCAGATTCCACAGAGTCTCTGGGCTTTTATTGCGACCAAAATCAGTGGAGAG 1830
Db 1897 aagcaatcagatctccaagatgcctctggccttattatgcgaccaaatccagtgagag 1956
QY 1831 TTCAATGTGTCAAGTCAATAGCAGAGATCAGTCAAGTCAAGAAAGATGAAGAGACAG 1890
Db 1957 ttcaatgtgtcagtaacaaatagcagagatcacctcgtgacaaagaagtaagagagcgg 2016
QY 1891 TGTTAGGGGGCAGAGAAATCAAAAGGGTCTTGGAAACCAAGGTCATAAAAATTACT 1950
|||||
Db 2017 tgttagggggcagagaaatcaaaagggtcctcttgaaagcaaaagtcataaaaaattact 2076
QY 1951 GCAGTTACTTACTCTGTCTCTGATGACCGGGGTCAATTCTCTCTTGACCAACTCCCCCT 2010
|||||
Db 2077 gcagttacttacctgtcttcttgatgacccgggtccattctctcttgaccaactccccct 2136
QY 2011 AGATTCAAGTGTAAAGAACTTCTGTTAGTGTCAACAGACCCCTCGAGTCTCTCTC 2070
Db 2137 agattcaagatgtaaagaatctctctgttagtgttcaacagccctctctgagttctctctc 2196
QY 2071 TACATCTGAGAGATATCTCTACATCCAATATGCATGGGTCACTGTTACAAGAGAACA 2130
|||||
Db 2197 tacatctgagagatatactcttaccatccaatatgcatggtgtccatcgttacaagagaaga 2256
QY 2131 CCGGATTTTGCAACAGTTCTCTGAGAAATGGGAATTCACCACTGAGAGTACCAAAATTAC 2190
|||||
Db 2257 ccggaatttgacaagatgtctgcagaaatggaattccaccagcttgaggtgccaagaattac 2316
QY 2191 TGCACAAAGCCATGGGAAAGACACCGAGTATTAATCTTGTGGGGAGGAAATGTTGT 2250
Db 2317 tgcagaagccacctgggaaagacacacagcagataaactctctgtggggacggaaatgtgt 2376
QY 2251 CAAGCAGAGCAGCTTAAGTCTTAAGAAGAAGAGATAATGCACTTCTTAGATACCTGCT 2310
Db 2377 caagcagagcagcctaagctcctaagaagaaggagataatgcaatctctgatactctgt 2436
QY 2311 GCACAGGATGATCTCTAGTATGCACTCTCTTAAGAACTACGCCCCAAGTGAAGAGCT 2370
Db 2437 ggcacagggaatgataccatgatagcactctctaagaactacacagccccaagtgaagaagtc 2496
QY 2371 GGCATAATAAATGAGTCACTGCAACCGATCCACACTTCTTACTCTCAAGTCAAGAGAAAGA 2430
Db 2497 ggaataaataaagatgacatgcacacagctccacacttccatagctcaagcaagaagaaga 2556
QY 2431 CCTAAATTAAGACAGACAACTAAGTGAAGAGGATCTGGAGACTTGGATTAATCTAGATGC 2490
Db 2557 ccttaaatatgaagaacagagaacagtgaaaggatcttgagacttgataactatgatagc 2616
QY 2491 TATTTCTGTTGATCTGACTAGTCTTCTGACTTTTACAATAATTCATATCTCAATAGTAC 2550
Db 2617 tatctctgttgatctgactagcttgcacttcttacaataatctccataatccccaatcgttag 2676
QY 2551 TCATCTGGGGATAGGCAACAGTGTTCATAGAACTAATTCCTGCGTTTAAANAAGTTC 2610
Db 2677 tcatcgggagactaagcaacagtgcttcaagaactaaatctctcgttggttgaagaatgc 2736
QY 2611 ACAGTCTGTCAGTCAATCTGCTCCATCATATTAACGAGCAGTGTCTGAGATAGCCCTGT 2670
Db 2737 acagctctgacagctatctcgtcccatataacagagcagtgctctgagataagccctgt 2796
QY 2671 TTCTGTGGCTCAAGTCTCTCAGTAAANAATATCAGTGTCTTCCCATGTTACCAAGACA 2730
Db 2797 tctcgttgctcaagctcctccagtaaaaaataatcaagtgtcttccccaatgttaccaaagca 2856
QY 2731 ACCCATGTTGGGTGGGAATCCAAAGATGATGATAGTACGAAATTTATGGTCAAGTAT 2790
Db 2857 acccatgtgtgggtgggaatccaaagaatgatagtatgtagtcagaaatatctgtgtcagaat 2916
QY 2791 GGGTGGGCGCAACCGAAATGTGACTGTGACTCAGACTCCTTCCTCAGTCAAGTGGGGCT 2850
Db 2917 g-----ggaagactgggctt 2931
QY 2851 ACCAAATCTMAAAGCGCGCAGATGAGAACCTATGATTCAAATCCATGGGAAGCACAG 2910
Db 2932 accaaatcraaaggccggcagaaatgaaacctatgaaatccaactccatcagtggaagccagg 2991
QY 2911 AGGAGATTTAATACTTCTTTACCAAGACTCTGAGTGGTGTCTATTCCACATTGCC 2970
Db 2992 aggaagattataacttcttcccaagacctgacatgggtgtgctctatctccacaatgccc 3051
QY 2971 TCTTGCTTAATAGATATCCAGTGCAGGTCGAGCAGTATTTGCAACGACGACAGATGCT 3030
Db 3052 tcttcggcttaataacataccacagtgcgagaccagatatactgcaaacagcagcaagatgct 3111
QY 3031 TCAAAATGAGGCTGTGTAATCCCATGAGGAAATGGGGCTTAATTCCTATGGCCAAGCAGC 3090
```


|||||
Db 397 aaaggaacaaagtaagacagatacagtcacaaataaagaagcaaggaaacacattcccaatga 456
OY 361 TGATGATGTTCAAAAAGCCGATGTATCTTCTACAGGCGAGGAGTTATTGATAAGACTC 420
Db 457 tgaatgagtcaaaaaagccgatgtatctctcaagggcagagtgattgataaagacc 516
OY 421 CTTAGAGCCGCTTTTACTTCAGGCAATGGATGGTTCCATTGTTGGTGAATCGACAGCC 480
Db 517 cttaaggaaccgcttttaactcaagcatlgatgttccctatttgylglaacgcagacg 576
OY 481 AAACATTTGATTGTATCGAAAAATGTCAACAATACCGCAATATTAACCAAGAGACCT 540
Db 577 aaacatgtatcttgtaacgaacaaagtcacacaaataccgcaataaagcaagagacct 636
OY 541 GCTTAACACAAAGTGTTTACATATCTTACATGAGAGAGACAGAAAAGATTTTCTTAAGAA 600
Db 637 ggttaacacaaagtglttaataatacttaacatgaagaagacagaaagatcttcttaagaa 696
OY 601 TTTTACCAAAATCTACAGTTAATGAGATTTCCTGACAAATGAGCCCCAAAGACAAAAAG 660
Db 697 tttaacaaatctacagtttaacatlgaggttccctgacaaatgagacccaagaacaaaag 756
OY 661 GCATACATTATTAATGCCGATGTGTGATGAACACACACATGATATTCTGGAAGACATAA 720
Db 757 ccatacatltaactgcgcgcatgltgatagaacacacacatgatalctggaagaacataa 816
OY 721 CGCCAGCTCTGAAATGCGCGACAGATATGAACAATGACAGTCTTTGCCCTCTCAGCC 780
Db 817 cgcagctctgaaatgagcgcagagatatagaacaaatgacgtgcttgcctgtctcaagc 876
OY 781 AGAGCTATGATGAGAGAGAGGGAAGATTGCAATCTTGTATGATCTGTGTGACGCCG 840
Db 877 aggaactaagatgaggaaggggaagatttgaatcttgtaagatctgtgtgacgcgcg 936
OY 841 CATTTACTACAGAGAAAGAAACATTTCCATCAAAACCTGAGACTTATTATCCAGACATGA 900
Db 937 ccttactacaggaagaacacattccacaaacccgagagcttataccaagaacatga 996
OY 901 TCTTTACAGAAAGTTGTCAATATGATACAAATTCACTGAGATCCTCCATGAGGCCCTG 960
Db 997 tcttccagaagaagtgltcaataatagatacaaatccatcgagatctctcaatgaagccgtg 1056
OY 961 CTTTGAAGATTAATCCGGAAGTGTATTGAGAGATTTTGTACTTAAATGATGAGGCAGTC 1020
Db 1057 ctltgaagataataccggaagtgltatcagaagatttltaglttaaaatgaaagcgagtc 1116
OY 1021 ATGTGCCAGAAAGCTCACTATCAAGA-----AGC 1050
Db 1117 atgtgccagaagaagcttactatacaagaagttaaccagtgatlggatacttccccaacagc 1176
OY 1051 TTAATCTTAATGGCATGACGAACCCGAGTATATGATTTCTGTTGGCTGATGAACTAT 1110
Db 1177 ttaacttaatgacatgacgaaccccgatatacgatctcgltgagtcgatacgagactat 1236
OY 1111 AGTGACTGCACAGCAAAAAGCAAACTCTTCGGAATCTGTAAACAAATGATCGACATGG 1170
Db 1237 agtgactgcacagcaaaaaaagaactcttccgaatactctgaacaaaatgaaatgcagatlg 1296
OY 1171 CTTTCTCTCAACCCACTTCTCTCAGAGAAACAGAAATGATATAGCAAAACCCAAATCC 1230
Db 1297 ctttgtctcaaacacacttctctcagaagaagaaacagaaatgaataabagacccaacatcc 1356
OY 1231 TGTTCGACAAAGGATTAAGACACCTATGAGCTGGATGCAACAGTTGCGTAGGCGCATGAG 1290
Db 1357 tgttcgacaaggaattagaccacactatgctgatacgacaacagttcgtagagcgcatlgag 1416
OY 1291 TATGTCGCAAAACCAAGCTTACAGATGCCGACAGCAGAGGCGCTATGGCTTGACAGACC 1350
Db 1417 tatgttcgcaaaacaaaggttacaagatgcgagcagcaaggtcctatggtcttgacagacc 1476
OY 1351 TAGCACACAGGAGAGATGAGTGAAGCTGATGAGGGGTTCAGTAACATGCTTCATT 1410
|||||

Db 1477 tagcaaccaagggcagatgagtggagctaggtataggggttccagtaacatgcttcatc 1536
OY 1411 GACCCCTGGGCGAGGCATGCAATTCACATCTTCTACAGAAACCAACATATAGGCTCAA 1470
Db 1537 gacccctgggcagagcatgcaataacacatcttccacgaacacaaactatagctcaaa 1596
OY 1471 CATGAGTACCCCCCAGATGGAGATCTGTGTTGTGCCCAAAACGACAGAAATATCATGAT 1530
Db 1597 catgagtaagcccccaacatgaggatcctgttcttccccaaacagcagaatataatgat 1656
OY 1531 TTTCTCTGTAATCGGGAGTCCAAAGATACCTTCACATCGATTTTTCCTGTTGACGG 1590
Db 1657 ttctccctgttaatcgvggagtcacaaagataagcctcaacatgatttctccgttgcagg 1716
OY 1591 TGTGCACTCTCCATGCGATCTTCTGTGCAATAGTGGGAACCAACAGCTTTTCAGCAGCTC 1650
Db 1717 tgtgcaccttcccatggaacttctctggaataactciggaaacacagcttctccaagcttc 1776
OY 1651 TCTCAGTGCCTTGCAAGCCATCAGTGAAGTGTGGGACTTCCTTTTATCTACTGTC 1710
Db 1777 tctcagtgccctgcaagccatcagtgaaagtgtgggagcttcccttatactactctgtc 1836
OY 1711 ATCACCGGCCCCAATTGGAATGATCTTCCATATGAAATATTACCCAAACCAAGTAAGT 1770
Db 1837 atcacagggcccaaatlgataaacttcccaataatgaatatlacccaacaaatgaagt 1896
OY 1771 AAGCAATCAGATTCCAAGAGTCTCTGGGCTTTTATTGCGACCAAAATCCAGTGGAGAG 1830
Db 1897 aagcaatcagatctccaagaagtcctctgggtcttcaatgtcgacaaataccagtgagagc 1956
OY 1831 TTCAATGTCTCAGTCAAAATAGCAGAGATCAGTCACTGACAAAGAAAGTAAGGACGAG 1890
Db 1957 ttcaatgtgtcagltcaaatagaacagatcacctcagtgacaagaagaatlaagaaagacag 2016
OY 1891 TGTGAGGGGGGAGAAATCAAGGGGTCCTTGGAAAGCAAGGTCATTAATAATTTACT 1950
Db 2017 tgttgaaggggcagaagaatcaaaaggggtcccttctgaaagcaaaaggtlcaataaaatctac 2076
OY 1951 GCAGTTACTTACCTGTCTTCTGATGACCGGGGTCAATTCCTCTTGACCAACTCCGCCCT 2010
Db 2077 gcagttacttaccgttctctctgtaagccgggttcaatctctctctgacccaactcccccct 2136
OY 2011 AGATTCAATTTTAAAGAAATCTTCTGTGATGTCACACAGCCCCCTCGAGTCTCTCTC 2070
Db 2137 agattcaagltgtaagaatctctctgttagtgtaaccagccccctgtgagttctctctc 2196
OY 2071 TACATCTGAGAGAGATCCCTTACATCCAAATGATGGGTGACTGTTAACAAGAGAGCA 2130
Db 2197 tacaatctggaagagatatactcttaacatccaataatgcaatggttcaatcgttaacaagaagca 2256
OY 2131 CCGGATTTTGCACAAGTTGCTGCAGAAATGGAAATTCACCAAGCTGAGTACCCAAAGATTAC 2190
Db 2257 ccggaatttgcacaagttgtctgcagaatggaattccaacagctgtagccaagaattac 2316
OY 2191 TGCACAAGCACTGGGAAAGACACACAGATTAATCTTGTGGGAGCGGAAATGTTGT 2250
Db 2317 tgcagaagcactatgggaagacacacagcagataaactctctgtggyagcaggaatgctgt 2376
OY 2251 CAAGCAGAGCAGCTAAAGCTTAAGAAAGAGAGATAATGCACTTCTTATATACCTGCT 2310
Db 2377 caagcagagagcagctaaagttcccaagaagaagagaataatgcaatctcttaagataactgtc 2436
OY 2311 GGACAGGATGATCTTATGATGCACTCTCTTAAGAACTACAGCCCCCAAGTGAAGAGAT 2370
Db 2437 ggaacagggatgatacttaagatgtaactcttaagaatacagcccccaagtgaagaggt 2496
OY 2371 GGACATAATAATGAGATCAGTGCACCAAGCTCCACCAATCTCTAGCTAAGTCAAGAGAAAGA 2430
Db 2497 ggaataataaataagatcagtgacacagctccacacatctcttagctcaagltcaagaagaaga 2556
OY 2431 CCGTAAATTAAGACAGAGACAAAGTGAAGAGGATCTGAGAGCTTGGAATCTTGATAGC 2490
Db 2557 cccataaatlaagaacagagaagaagtgaaagggatctggaagacttgatataatctatagtc 2616
|||||

QY	2491	TATTCCTGGTATCTGACTAGTTCCTACATTTTTCATTAATTTCATATCTCTCAAAATGGTAG	2550
Db	2617	tattcttgygactcgtactagcttcttgcactttaccataaattccatatctccaaatgyttag	2676
QY	2551	TCATCTGGGGACCTAACCAAGAGTGTTTCAGAGCACTAATTCCTGGGTTTGAAAGATTC	2610
Db	2677	tcaatcgyggactcaagaacaagtgcttcaaggaactaaattccgggttcgaaagttc	2736
QY	2611	ACAGTCTGTGACAGTCAATTCGTCTCCATATACCGACGACGTCTCTGGATAGCCCTGT	2670
Db	2737	acagtcctgygactatctctgcctccatataaacggagagtgctctcgtatagccctgt	2796
QY	2671	TTCTGTTGGCTCAAGTCTCTCAGTAAAAAATATCAGTCTTTCCCATGTTTACCAAGCA	2730
Db	2797	ttctgcttgyctcaagttccctccagtaaaaaatacagtgcttccccaatglttaccaaagca	2856
QY	2731	ACCATGTTGGGTGGGAATCCAGATGATGGATGTCGAGAAATTTATGGCTCAAGAT	2790
Db	2857	accacatgcttgyggttgggaatcccaagaatgtagtgcagaaataatgagctccaat	2916
QY	2791	GGGTGGGCCAAACCGAATATGACTGTGACTAGACTCTTCCTCAGAGACTGGGCTT	2850
Db	2917	g-----gagagactgyggtc	2931
QY	2851	ACCAACTCAAAAGCCGGCAGAAATGGAACCTATGAAATTCAACTCCATGGGAAGCCAGG	2910
Db	2932	accaactcaaaagccggcsgaaatggaactatgaatccaactccatggaagacag	2991
QY	2911	AGGAGATTTATATATCTTCTTACCAGACTGTGACGTGGGTGGCTATTTCCACATTGCC	2970
Db	2992	aggagatataaatacttcttaccacgacctgacatggttgcatacttccacaatgctc	3051
QY	2971	TCCTCGGCTAATATATGCTATACAGCTGCGAGACAGATTTGCAAGACGACAGACATCT	3030
Db	3052	tcttcggtcctaataagctaccacagtgycgaacacagatatgcaacagccagccagatgctc	3111
QY	3031	TCAAATGAGGCTGTGTGAAATCCCATGGGAATGGGGCTAATCCCTATGGCCAGCAGC	3090
Db	3112	tcaaatgagggcttgytgaatcccccataggaaatgggggttaatccctatgyccaagcagc	3171
QY	3091	AGCATCTAACCACTGGGTTCTCTGGCCCATGGCATGTTGTCCATGGAACAAGTTTCTCA	3150
Db	3172	agcatctaaaccaacttgygtcttcctggcccgatgycatgctgtccatgycgaacaagtgttcca	3231
QY	3151	TGGACATCAAAATFAGGCTCTCTCTCTTGAAGAAATCCCTGATGATCTGTTGGCCACCTTC	3210
Db	3232	tgycaactcaaaaatagggctctcttcttctttagaatctcccgcgtatcttgytgggccaacttc	3291
QY	3211	CAACCTGGAAAGCCAGATGACGAAAGAGAGCATTTATGGACGACGCTGACACTCTTCTTAG	3270
Db	3292	caaccttgyaagggccagaggtgaacgaagaagatatttgygccaagctgcgaacactcttccag	3351
QY	3271	CAACACAGATGCGACAGGCTGTGAAGAATTGACAGAGCTTTGGGCATTTCTGAACTTGT	3330
Db	3352	caacaacagatgcccacagagccttgyaagaanaattgacagagcttcttgycatcttcgtactgtt	3411
QY	3331	CAATCAGGAGACAGCATTAAGACCCCAAACAGATGCTTTTCCAAGGCCAAGAAAGCAGCAGT	3390
Db	3412	caatcagtggaacaggtcatatagacccaaaacagatgtcttccaaaggccaaagaagcagcagt	3471
QY	3391	AATGATGGATCAGAAAGCAGAGATTAATATGACAGCATCCAGACAGAGGGCCCTCCAT	3450
Db	3472	aatgatgtatctagaagggccagatataatgtaacagaaataccacagacaggggcccacat	3531
QY	3451	GCAAGAGGCTTTTCATCTTGAGGAGCAATCCACATCTTTAACTGTATGATGAATCAGAT	3510
Db	3532	gcaagggaggtctcatcttccagggacaatcacatctttaactctatgataatcagat	3591
QY	3511	GAACACGACAGGCAATTTTCTCTCCAAAGAAATGCAACCCACAGACCAACATCATGAGCC	3570
Db	3592	gaaccagcaaggtcaattctcttccaaaggaatgcaacccacagacgaacatatcatgagcc	3651

[illegible]

```

RESULT 1455-11058
US-09-652-355-11058
; Sequence 11058, Application US/09652355
; GENERAL INFORMATION
; APPLICANT: Shytan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEOTIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1192-001

```

; CURRENT APPLICATION NUMBER: US/09/652,355
 ; CURRENT FILING DATE: 2000-08-30
 ; PRIOR APPLICATION NUMBER: 60/151,136
 ; PRIOR FILING DATE: 1999-08-30
 ; NUMBER OF SEQ ID NOS: 11227
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 11058
 ; LENGTH: 7888
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-652-355-11058

Query Match 95.9%; Score 4313; DB 25; Length 7888;
 Best Local Similarity 97.9%; Pred. NO. 0;
 Matches 4436; Conservative 0; Mismatches 10; Indels 87; Gaps 3;

QY 1 GGTGATGTTGACACGACGACCAATTAATAAATGCTGTTGACATCTCTTGACGCTGTT 60
 DB 97 gctgagatggtgactcagacccaataaataaactcgtcgaacatccttgactggtt 156
 QY 61 ACCGAGTTGCTGATGATATTCAGATGAGTATGAGAGAAACTTGATGCACTGCGC 120
 DB 157 agccagctgctgactgatactcagaatgagatgagatgagaaacttgatccactgyc 216
 QY 121 CAGTGATTCAGCAAAAGCAAAATGCGCATGTGATCTCCAGACAAAGTCTTACCTGAC 180
 DB 217 cagtgatcagcaaaacgaacaaatgcacatgatactccagacaaagctcctacccgag 276
 QY 181 TGGTGAAGAAAGGAG 240
 DB 277 tgggtgagaaag 336
 QY 241 ATCTGCAATCTTATGATATGACATTTTCAATGTCACCAAGATTAATGTCGATTTT 300
 DB 337 atctgccaatctttagtgatatactgacaatttcaatgacaaacgaataaactgagatctt 396
 QY 301 AAGGAGACAGTAAGACAGATACGTCATTAATAAAGCAAGAAATATTTCCATTA 360
 DB 397 aagagaaacagtaagacagatacgtcaaaataaagagcaagaaacatctccatgaa 456
 QY 361 TGATGATGTTCAAAAAGCCGATGATCTCTACAGGGGAGGGAGTTATGATTAAGACTC 420
 DB 457 tgatgagatgtaaaaaagccgaatgatactctcacaagagagagatgataaataacc 516
 QY 421 CTTAGAGACGCTTTTACTTACAGGATGAGTGGTCTTCTATTTGTTGTTGATGAGAGCC 480
 DB 517 cttagagaccgcttttactctcagcatgagatgagtttccctatcttggtgagatcgagag 576
 QY 481 AATCATTTGATTTGTTGTAAGAAATGTCACATATCTGCAATATTAACAGAGAGACT 540
 DB 577 aatcatctgatttctgtaacgaatgtaacacatcctgcaataataaagaagaagacct 636
 QY 541 GGTTAACAAGATGTTTACATATCTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 DB 637 ggttaacaagatggtttacataatcttacaatgaagaagaagaagaagatcttccataaga 696
 QY 601 TTTTACCAAAATCTACAGTTAATGAGATTCTCTGACAAATGAGCCCAAGACAAAAAG 660
 DB 697 tttaaccaaaatctacagtttaatgagattctcctgacaaatgagacccaagaacaaaaag 756
 QY 661 CCATACATTTAATGCGCGATGTTGATGTAAGAAACACACATGATATTTCTGGAAGACTAA 720
 DB 757 ccatacatttaatgcccgaatgtaagtaaaacacacataatcttgaaagacataaa 816
 QY 721 CCGCAGTCCTGAATGCGCGAGAGATATGAACAAATGCAAGTCTTGGCCTGTCAGGC 780
 DB 817 ccgcagtcctgaatgcccgaagatatagaacaaatgcaagctcttggcccttctcaagcc 876
 QY 781 ACGACCTATGATGAG 840
 DB 877 acgagctatgag 936

QY 841 CATTAACAG 900
 DB 937 cacttaacag 996
 QY 901 TCTTCAG 960
 DB 997 tcttcag 1056
 QY 961 CTTTGAGATTAATTAACAG 1020
 DB 1057 ctttgag 1116
 QY 1021 ATGTCGCCAGAAAGCTACTATCAAGA-----AGC 1050
 DB 1117 atgtccccaagaacgctactatcaacagagagagagagagagagagagagagagagagag 1176
 QY 1051 TTAATCTTAATGAG 1110
 DB 1177 taatccttaag 1236
 QY 1111 AGTGAATGAG 1170
 DB 1237 agtgag 1296
 QY 1171 CTTTGTCTTAACCCACTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230
 DB 1297 ctttgtcttaacccacttcttcaagagagagagagagagagagagagagagagagagag 1356
 QY 1231 TGGTGAAG 1290
 DB 1357 tgggtgag 1416
 QY 1291 TATGTGCCAAACCAAGGCTTACAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1350
 DB 1417 tatgtgccaaaccaaagcttacagatgagagagagagagagagagagagagagagagag 1476
 QY 1351 TACGACACAG 1410
 DB 1477 tagcacacag 1536
 QY 1411 GACCCCTGGGAG 1470
 DB 1537 gacccctgggag 1596
 QY 1471 CATGAGTAGCCGCCACATGGAAGTCTGAGTCTGCCCCCAACAGAGAGAGAGAGAGAG 1530
 DB 1597 catgagtagccgccacaatggagatctgagcttcccccaacagagagagagagagag 1656
 QY 1531 TTTCTCTGTAATGCTGGAG 1590
 DB 1657 ttctctctgtaatgctggag 1716
 QY 1591 TGTGACACTTCCCATGAG 1650
 DB 1717 tgtgacacttcccatgag 1776
 QY 1651 TCTCAGTCCCTGCAAGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1710
 DB 1777 tctcagtcctctgcaagccatgagagagagagagagagagagagagagagagagagagag 1836
 QY 1711 ATCACCAGGCCCAATTTGATGATCTTCCCAATTAATGATTAATTAACCAACCAAGTAAG 1770
 DB 1837 atcaccaggcccaatttgatgattcttcccaatttaattgatttaacccaacaaagtaagt 1896
 QY 1771 AAGCAATCAGATTAATCAAG 1830
 DB 1897 aagcaatcagatttaacag 1956
 QY 1831 TTTCAATGTGTGAGTCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1890
 DB 1957 ttcaatgtgtgagtcaaaatgag 2016
 QY 1891 TGTGAGGGGAG 1950

Qy	4099	GATCGAACACCCCGAGGCTGCATTCATCTATCATGTCCTCGAATAATGAAGGCTGGCCATC	4158
Db	4192	gattcaacaccccgaggctctcatctcatctacagtcctctcagaatgaaggagcttgccatc	4251
Qy	4159	AGGAAATTTGGCCGAGGAACAGCTCTTTTCCGACGACGTTTGGCCACACAGGGGAATCC	4218
Db	4232	aggaaatcttgccaggaaacagctccctctcccgacgacgatttgcacacaggggaaacc	4311
Qy	4219	TGCAGTGTATAGTATGCTGCACATGATGGCAGCACTGGTCACATGGGACAGATGAACAT	4278
Db	4312	tgcaagtataagtatggttgacaatgaatgagcagctgtctacatcgggaacagatgaacat	4371
Qy	4279	GAACCCCATGCCATCTGCTGGCATGECCTATGGTCTCTATCAAGAAATACCTGTCACATCT	4338
Db	4372	gaaccccatgcccatctgcttgccatgctcctatggctcttgatcgaatctactctgacatct	4431
Qy	4339	CTGCACACGAGACGCTTAAGGAACACAGTGCACATGACATGACATGAAATGATTTATTTGG	4398
Db	4432	ctgcaacaggaacctcttaaggaacaacgtaacaaatgacatctgacactaaggatattggg	4491
Qy	4399	AAGGAATCATTTGTTCCAGGAGCATCATCTTGGAAGAAAGACAGCTTTGAGCTCCATAA	4458
Db	4492	aaggaaatcatctgttcacgagatcatcattctggaagaaggacagcttgagctcatcaaa	4551
Qy	4459	GGGTATTTTAAGTGTATCTCAATTTGACGAGAAAT	4491
Db	4552	gggtatcttaagtgatgtcatctttggaacaggaact	4584

```

RESULT 15
US-09-663-1028
; Sequence 1028, Application US/09663693
; GENERAL INFORMATION:
; APPLICANT: Gearling, David P.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1195-001
; CURRENT APPLICATION NUMBER: US/09/663,693
; CURRENT FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/154,986
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 1340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1028
; LENGTH: 7888
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-663-693-1028

```

Query Match	95.98;	Score 4313;	DB 26;	Length 7888;
Best Local Similarity	97.98;	Pred. No. 0;		
Matches 4436; Conservative	0;	Mismatches	10;	Indels 87; Gaps 3

QY	1	GC	GGATGCTGCTACTCAGAGACCAATAAATAACGCTTTGAACATCCCTTTGACTGCTT	60
Db	97	gc	tgaagagggagagaccagaagaccataataaacgcgtctggaacaccccttcgacggtc	156
QY	61	AG	CCAGTGTGCTGATGTATATTCAAGATGACTGGATTGAGAAACCTTGATCCACTGGC	120
Db	157	ag	ccagcctctgcgacgctatctccaagatcgagctgagcttagagaaaccttgaatccactgyc	216
QY	121	CAG	TATTCACGAAACCAATAATGCCATGTGATATCCCAAGGACAACGCTTTACCTGCAG	180
Db	217	cag	tatctcaagaaaacgcaaatctgccaatgatactccagagcaagggctcttaacctgcag	276
QY	181	TG	GTGAAAAACGGAGACGGGAGCGAGGAAAGTAAATATATTGAACATTTGCTGACTGAT	240
Db	277	tgg	tgaaaaaagagagacgaggaagctaaatatatctgaagaaatctggcctgaagctgat	336
QY	241	AT	TGCCAATCTTGGTGTATATTGCAATTTCAAAAGTCAACCAAGATTTAAAGTCCGATTTT	300

Dd	337	atcgcgcaactcttagtgatataatgacaaattccaatgctccaacccagataaagtctgcgaattt	336
Qy	301	AAAGGAAACAGTAAAGACAGATACGTCGAATMAAGAGCAAGGAAAACTATTTCCATATGA	360
Dd	397	aaaggaaacagtaagacagatacgtccaataaagaagcaaggaataactattccaatga	456
Qy	361	TGATGATGTTTAAAAAGCCGATGATCTTCTACAGGCGAGGAGATTATTGATMAAGACTC	420
Dd	457	tgatgatlctcaaaaagccgatlcatctctcaagggcagggagtatctgataaagactc	516
Qy	421	CTTAGGACCGTTTTACTTCAGGAGATTTGGATGGTTTTCTATTGGTGGTAATCGAGAGGC	480
Dd	517	cttaggacgccttctaactccaaggaaattggaatggtcttccattcttggtgtaacgagacg	576
Qy	481	AAACATTGTATTTGTATCAGAAATGTGCACACATATCTCGCAATATMAAGCAAGAGACCT	540
Dd	577	aaacctgtattctgtatccagaaatgtcacacaataacctgcgaataaagcaagggacct	636
Qy	541	GGTTAACACAGTGTTTACATATCTTACATGAAGAAACACAGAAAGATTCTTAAAGA	600
Dd	637	ggttaacacaagtgcttaacaatacttacaatgaagaacagaaagatcttcttaagaa	696
Qy	601	TTTACCAAAATCTACAGTTTATGAGAGTTTCTCGACAAATGAGCCCCAANAGCAAAAAG	660
Dd	697	cttaccaaaactctacaglttaatggagttccctcggacaaaatgaaacccaagaacaaaag	756
Qy	661	CCATACATATTTAATTCGCCGTATGTTGATGAAGAAACACACATGATATTCTGGAAGACATAA	720
Dd	757	ccatcatatttaattgcggtatgttgatgaataacacaaatgatatctctgaaagacataaa	816
Qy	721	CGCCAGTCTTAAATGCGCCAGAGATATGAAACAAATGCAGTGTCTTGGCCCTGTCAGCC	780
Dd	817	cgcagatctctaaatctgcgccaagataatgaacaaatgcagtgctctgcctctctccacgc	876
Qy	781	ACGAGCTATGATGAGAGAAAGGGGAAAGTTTGCATCTGTATGATCATGCTGTGTCACGCGG	840
Dd	877	acgagctatgataaggaagaaggaaatcttgcgaactctgataatctctgtggtgcgcgcg	936
Qy	841	CATTACTCAGAGAAAGAACATTTCTCANTCAAAACCTGAGAGCTTTATACACAGCATGA	900
Dd	937	catctactacaggaagaagaacattccaatcaaaacctgagagcttatattacaagacatga	996
Qy	901	TCTTTACAGAAAGTTGTCAATATAGATACAAATTCATGAGATCCTCATGAGCCCTGG	960
Dd	997	tcttccaagaaaaggtctgcaataatagaatacaaatctacatgagctctccatagaagcctgg	1056
Qy	961	CTTTGAGATATTAATCCGAGAGGTATTCAGACATTTTATAGCTAAAGATGGGCATGC	1020
Dd	1057	cttgaagaataataactccgaaggtgtatactcagaattctttagcttaaaatgaatggtgcagtc	1116
Qy	1021	ATGGTCCGACAAAGCTACATATCAGA-----AGC	1050
Dd	1117	atggttccagaagaatgtacatactcaagaagtctaccagtgatgtagatacttctcccaaacgc	1176
Qy	1051	TTATCTTAATGCGCATGCAGAAACCCGATATATGCAATTTCTCGTTGGCTGATGGAATAT	1110
Dd	1177	ttatcttaatgccaatgcagaaacccagatatctgattctcgtctgctcgtatgaactat	1236
Qy	1111	AGTAGCTGCACAGACAAAAAGCAAACTCTTCCGAAATCTCTGTATACAAATGATGCACTATG	1170
Dd	1237	agtactatgcacagacaaaagaacaaactcttcgaaatctctgtatacaaaatgatactgcagatcg	1296
Qy	1171	CTTTGTCTCAACCCACTTCTCTCAGAGAGAAAGATGGATATATACGCCAAACCCAAATCC	1230
Dd	1297	ctctgtctccaacccaattcctctcagagaagaacagaaatgataaagccaacccaactcc	1356
Qy	1231	TGTTTGACAAAGGATTTAGACCACCTATGAGCTGATCAACAGTTGGTATAGCGCGCATATG	1290
Dd	1357	tgttggaacaaggatctagacaaactatagcttgataatgcaacagtttcggtaggtgcgaatag	1416
Qy	1291	TATGTGCGCAAAACCAAGGCTTACAGATGCCGAGACAGAGGCGCTATATGCTTGGCAGACCC	1350
Dd	1417	tatgtgcgcaaaacaaagcttacaagaatgcgcaacaaacaaagcctatagctcttgcgtgaagcc	1476

QY	1351	TTGGACACAGGGGAGATGAGTGGAGCTGATGGGTTCCAGTAACATACCTTCATT	1410
Db	1477	tagaccacaaggccagatgagatgagactagatagtagtgaggggttccagtaacatagcttcatt	1536
QY	1411	GACCCCTGGGCGAGCATCATCATCTTCCTACAGAACCAACATATGGGCTCAA	1470
Db	1537	gacccctggccagagcatgacatcaactccttccacgaacaaactatagctcaaa	1596
QY	1471	CATGATAGCCCCCACATAGGAGATCTGTGCTTGCCCAACACAGACAGATATTCATGAT	1530
Db	1597	catgtagccccccaacatgaggatcctgtcttgcaccaaccagcagaatatcatgat	1656
QY	1531	TTTCTCTGCTAATCGTGGGAGATGCCAAGATAGCCTTCACATCAGTTTTCCTCTGTGCAGG	1590
Db	1657	ttctctctgtaatcgttggagatccaaagataagacctacacatcagtttctctctgttgcagg	1716
QY	1591	TGTGCACTTCCCATGGGCACTTCTTGCGCAATACTGGGAACACAGCTTTTCCAGCAGCTC	1650
Db	1717	tgtgcacttcccataggacatctcttctgtgcaatacttggaaaccacagtttctccagcagctc	1776
QY	1651	TCTCAGTGCCTTCGAAGCCATCAGTGAAGGTGTGGGACTTCCCTTTATCTACTGTCTC	1710
Db	1777	tctcagtgccctgcgaagcccatcagtgaaagtgtygggagacttccctttatctactctgtc	1836
QY	1711	ATCACAGGCCCCCAATTTGGATTAACCTCCCAATATGATATATACCAACCAAGTAAAGT	1770
Db	1837	atcacaggccccaaatltgataaactctcccaatatgaatatataccaaaccaagtaaaagt	1896
QY	1771	AAGCATCAGATTCATCCAAAGAGTCTCTGGGCTTTTATTCGACCAAAATTCAGTGGAGAG	1830
Db	1897	aagcaatcaggatcccaagatccctctgtggtctttatctgtgaccacaatcccaagltgaaag	1956
QY	1831	TTCAATGTGTCACTCAATAGCAGAGATCACTCAGTGAACAAAGAAATPAGGAGACAG	1890
Db	1957	ttcaatgtgtcagtcataatagacagatcaactcagtgacaaagaaglaaagaagagcag	2016
QY	1891	TGTTAGAGGGGAGAGAAATCAAGGGGCTTTGGAAAGCAAAAGTCATPAAAAATTTACT	1950
Db	2017	tgttgaaggggagagaaatccaaggggtcctcttggaaagcaaaaggtccataaaaaattact	2076
QY	1951	GCAGTTACTTACCTGTTCTTCTGTGATGACGGGGCTATTCTCTCTTGACCAACTCCCCCT	2010
Db	2077	gcagttacttacctgtcttctctgtatgaccgggttcaatctctcttgaccaactccccct	2136
QY	2011	AGATCAAGTTTAAAGATCTTCTGTAGTGTACACAGCCCTCTGGAGTCTCTCTCTC	2070
Db	2137	agatccaagttaaagaatctctctgttagtgtcaccaagccccctcgtgagttctccctc	2196
QY	2071	TTCATCTGGAGAGATCCTTACATCATCATATGATGATGGGTCACTGTTACAGAGAACA	2130
Db	2197	tcatcttgaggagatcctctcaacatcaatagcatgtggtctcgttitaacaagaagaaca	2256
QY	2131	CCGGATTTTGGACAAGTTGCTCTGAGATGGGAATTCACACAGCTGAGTAGGCCAAGATTAC	2190
Db	2257	ccggattttggacaagttgtctgcagatlggaattccaccagctgagtagtcccaagtttac	2316
QY	2191	TGCACAAGGCATGGGAAGACACCGACGATTAATCTTGTGSGGAGGGAATTTGTG	2250
Db	2317	tgcacaagaccttgggaagacacccagcatlaaactctctgtcgggagcagaaatgttgt	2376
QY	2251	CAAGCAGAGCAGCTAAGTCTTAAGAAGAAGAGAAATATGCACTTCTTAGATACCTGTCT	2310
Db	2377	caagcagagcagctaaagcttaagaagaagagaaatagcatcttctaagataactgtct	2436
QY	2311	GGACAGGGATGATCTCTAGTAGTACACTCTCTTAAGAACTACAGCCCCCAAGTGGAGAAGT	2370
Db	2437	ggacaggagatgatactctaagtgatcactctctaagaactacagcccccaagtggaaagagt	2496
QY	2371	GGACATATAAAGAGTCACTGACACAGCTTCACCATTTCTAGTCAAGTCAAGAGAAGA	2430
Db	2497	ggataataaatagagtctagtgcaccagcttccaccattctctagtctcaagtccaagaagaaga	2556

QY	24.31	CCGAAATTAAGACAGACAGCATGTGAAGGGATCTGGAGCTTGGAATCTAGATGC	24.90
Db	25.57	cccttaaatltaagccagagacaagtgaaagggatctggagactggaatactagaagc	26.16
QY	24.91	TATCTTGTCATGTGACATGACTTCTGACATCTTTTCATTAATCTCATATCTCAATATGTAG	25.50
Db	26.17	tattcttgggatactgactgttccgaactttaaataatccatactccaaatggtag	26.76
QY	25.51	TCATCTGGGGCATTAAGCACAGGTGTTTCAAGGACATTAATCTCTGGGTTTGAAAAGTTC	26.10
Db	26.77	tcaatctgggactaagacaacagtglttcaaggaactaactctcgggtttgaaagtc	27.36
QY	26.11	ACAGCTGTGTCAGTCTATTGTGCTCTCATATAACGACAGTGTCTTGATAGCCCTGT	26.70
Db	27.37	acagctcgtgtagactctatctgctccataataaccgagcagtgctctcggataagccgtg	27.96
QY	26.71	TTCTGTTGGGCTCAAGTCTCTCAGTAAAAAATATACATGCTTTCCCATGTTTACCAAAACA	27.30
Db	27.97	ttctgcttggcccaagctccccaaglaaaaaatacaagtgcttcccagtttaacaaagca	28.56
QY	27.31	ACCCATGTTGGGTGGGAATCAAGAAATGATGATATCATCGAAAAATTAATGCTCAAGTAT	27.90
Db	28.57	acccatgttgggtgggagatcccaagaatgatgagtagtcaagaaatatatgctcaagtat	29.16
QY	27.91	GGGTGGGCCAAACCGAAATGTACTGTGACTGACTGACTCTTCTCAGAGACTGGGCGTT	28.50
Db	29.17	g-----g-----gagactgggctt	29.31
QY	28.51	ACCAAACTCAAGGCGGGCGGAATGGAACCTATGATTTCAACTCTCATATGGGAAGCCAGG	29.10
Db	29.32	accaaaactcaaaagccggccagaaatggaactaagaaatcaaatcccatgggaagccag	29.91
QY	29.11	AGGAGATTAATTAATCTTCTTACCACAGACTCGACATGGGTGCTCTATTCCACATTGACC	29.70
Db	29.92	agggagatataataactctcttaccacagactcgactcgggtgtctctatcccaattggc	30.51
QY	29.71	TCCTTGCTTAATATGATACCATCCAGGTGGAGACCAGTATTGCAACAGCAGCAGACAGTCT	30.30
Db	30.52	ttcttggtcctaattagcataccaggtgcgagaccagtatgtcaacagcagcagatgct	31.11
QY	30.31	TTCAATGAGGCTGGTGAATATGCCCATGGGAATGGGGGCTATTCCTTATGGCCAAAGCAGC	30.90
Db	31.12	tcaaatgggactcggatgaataatcccaatgggaaatggggcttaactccatctgccaagcagc	31.71
QY	30.91	AGCATCTAACCAATGTGGGTTCTGTGGCCGATGGCATGTTGTGCATATGAAACAAGTTTCTCA	31.50
Db	31.72	agcatctaaaccaactgtgttccggcccgatgcatggttccatggaacaagtctccca	32.31
QY	31.51	TGGCATCTCAAAATAGGCTCTCTTCTTAGGAATTCCTGTGATGATCTTGTGGCCACCTTC	32.10
Db	32.32	tggcactcaaatatagcctctctctcttagaatccctgtgatacttcttgggcccaccttc	32.91
QY	32.11	CAACTCGGAAGCCGACAGGTGACGAAGAAGAGATTAATTGGACACAGCTGACACTTTCTAG	32.70
Db	32.92	caactctggaagccaaggtgacgaagaagatattggaacagctgcaacactctctcag	33.51
QY	32.71	CAACACAGATCTCGACAGCCTTGGAAGAATTTGACAGATTTTGGCATTTCTCGAACTGTG	33.30
Db	33.52	caacaacagatgccaacagccttggaagaataatgcaacagatcttgggcatcttccgaactgt	34.11
QY	33.31	CAATTCAGGACAGGCATTTAGAGCCCAAAACAGATGCTTTCAGAGGCCAAGAAGCAGCAGT	33.90
Db	34.12	caatcaggaagacagcatctagagcccaaaacagatgtcttccaagccaagaagcagcag	34.71
QY	33.91	AATGATGATTCAGAGGACAGGATTAATATGACAGCATATCCACGACAGAGGGCCCTCCAA	34.50
Db	34.72	aatatgatatcaagaagccagatctatctatggaacagataccacagacaggggcccacat	35.31
QY	34.51	GCAAGAGAGGCTTATCTTACAGGACAAATCACCATTTTAACTCTATAGATGAATCAGAT	35.10
Db	35.32	gcaagggaggtcttaactctcaaggaacaatccatctttaaactctatgataagatcagat	35.91
QY	35.11	GAACCCAGCAAGCAATTTCTCTCTCCAGGAATGACCCACGAGCCAAATCATGAGCC	35.70

```

Db 3592 gaaccagcaggaatttccctcccaagaaatgcacccacagccacacatcagagcc 3651
      |||
QY 3571 CGGACAAACACCCCAAGCAATTAGATGAGCTTACAGAGGCTGACAGGCCAGCA 3630
      |||
Db 3652 ccggaacaaaccccccaagcaacttagaaatgcagcttcagcagagctgcagggccagca 3711
      |||
QY 3631 GTTTTGAATCAGAGCCGACAGCATTGAATTGAAATGAAACCTTACTGCTGGTGG 3690
      |||
Db 3712 gttcttgaatcagagccgaagccttgaatlgaaatlgaaacccctacgctgglyg 3771
      |||
QY 3691 TGCCTCGGTATGAGGCTTATGATGAGCCGACAGGCTTCTTATGCTCAATGCT 3750
      |||
Db 3772 tgcctcggtatgagggccttatgctcagcccaagcagggcttcttaatgctcaatgct 3831
      |||
QY 3751 CGCCCAACGACAGCAGAGCTGCTAAGTCATCTTCCGACACAGAGGCTGCTATGAT 3810
      |||
Db 3832 cgcacaaacagcagagagctgctgaatcactcgcgaacacagagggctgctatgat 3891
      |||
QY 3811 GATGACGACGAG-----CAACAGACGACGACGACGACGACGACGACGACGACGAC 3858
      |||
Db 3892 gatgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 3951
      |||
QY 3859 GCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAACAGCAAC 3918
      |||
Db 3952 gcaacagcaacagcaacagcaacagcaacagcaacagcaacagcaacagcaacagcaac 4011
      |||
QY 3919 TGCCTTCCCAAGATGAGTGGCTTTGGCAGGACCCACCAATGCCAAGCTCTCCGCA 3978
      |||
Db 4012 tgccttcccaagatgagtggtcttggcaggaaccacacatgcccacagctccctcgca 4071
      |||
QY 3979 ACAGTTTCCATATCAACCAATTTATGGAATGGGACACACCAAGATCCAGCTTGGTGG 4038
      |||
Db 4072 acagtttccatatacacaacaaatlatgaaatgggacacacacacagatccagccttgglyg 4131
      |||
QY 4039 AGTGTACTAGTCTCCCAATGCAATGATGTCGTCAGAAATGGTCCCTCCAGATCCCAT 4098
      |||
Db 4132 agtgtactagctccccaatgcaatgctgctcgaagaatgggctccccaatcccat 4191
      |||
QY 4099 GATGCAACACCCGACGAGCTGATCCATCTATCATCTTCCAGAAATGAAGGCTGGCCATC 4158
      |||
Db 4192 gatgcaacacccgacgagctgcatctcatcagctcagaaatgaaagctggccatc 4251
      |||
QY 4159 AGGAATTTGGCCAGGACAGCTCTTTTCCAGACAGAGTTGGCCACAGGGGATCC 4218
      |||
Db 4252 aggaatlttggccaggaacagctccttccacagcagcttggccacaggggaatcc 4311
      |||
QY 4219 TGCAGTGTATGATGTCACATGATGAATGAGCAGTGGTCACATGSGACATGAACAT 4278
      |||
Db 4312 tgcagtgtatgattggtgacatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 4371
      |||
QY 4279 GAACCCATGCCATGCTGCGATGCGTATGGCTCCTGATCAGAAATACCTGCTGACATCT 4338
      |||
Db 4372 gaacccatgcccattgctgcatgctctatggctctgctgctgctgctgctgctgctgct 4431
      |||
QY 4339 CTGCAACAGGACCTCTTAAGAAACACCTGTACAAATGACACTGCACTAGATTTATGGG 4398
      |||
Db 4432 ctgcaacaggaacctcttaagaaacacatgtaaaatgaaatgaaatgaaatgaaatgaaat 4491
      |||
QY 4399 AAGGAATCTTGTTCAGGACATCTTGAAGAAAGAGACAGCTTGAAGCTTGAAGCTTGAAG 4458
      |||
Db 4492 aaggaatcctgttccagcactcctggaagaagacagccttgaagcttccatcaa 4551
      |||
QY 4459 GGGATTTTAACTGATGTCATTTGAGCAGGAAT 4491
      |||
Db 4552 gggatlttaagtgtatgctattgagcaggaact 4584
      |||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2001, 17:30:22 ; Search time 99.89 Seconds
(without alignments)
7856.782 Million cell updates/sec

Title: US-09-041-994-1
Perfect score: 4496
Sequence: 1 GCTGATGCTGACTGACAGAG.....CATTTGACAGCAATTTCTAG 4496

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 302621 seqs, 87301344 residues

Total number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PTOTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.6	1.7	397	4	US-09-253-691-3
2	73.2	1.6	168	1	US-08-469-802B-4
3	73.2	1.6	168	2	US-08-267-803B-4
4	73.2	1.6	171	1	US-08-469-802B-5
5	73.2	1.6	171	2	US-08-267-803B-5
6	73.2	1.6	195	1	US-08-469-802B-2
7	73.2	1.6	195	2	US-08-267-803B-2
8	73.2	1.6	234	1	US-08-469-802B-3
9	73.2	1.6	234	2	US-08-267-803B-3
10	68.4	1.5	154	1	US-08-469-802B-6
11	68.4	1.5	154	2	US-08-267-803B-6
12	68.4	1.5	506	1	US-08-469-802B-7
13	68.4	1.5	506	2	US-08-267-803B-7
14	68.4	1.5	3366	1	US-08-469-802B-1
15	68.4	1.5	3366	2	US-08-267-803B-1
16	68.4	1.5	10660	2	US-08-267-803B-8
17	67.6	1.5	10348	3	US-08-457-273B-41
18	67.6	1.5	10348	3	US-08-556-419-13
19	67.6	1.5	10366	1	US-08-246-982A-5
20	67.6	1.5	10366	1	US-08-453-265-5
21	66.4	1.5	3376	1	US-08-320-559-29
22	66.4	1.5	3376	3	US-08-545-860D-29
23	66.4	1.5	3376	5	PCT-US94-04496-29
24	64.8	1.4	2770	4	US-09-008-697A-13
25	64	1.4	379	4	US-08-145-617-5
26	63.2	1.4	1776	2	US-08-531-927B-1
27	60.6	1.3	543	6	5273901-6

28	59.8	1.3	1884	3	US-08-753-007A-5	Sequence 5, Appl1
29	59.8	1.3	1884	4	US-09-398-496-5	Sequence 5, Appl1
30	59.6	1.3	325	2	US-08-531-927B-3	Sequence 3, Appl1
31	59.4	1.3	4835	1	US-08-386-495-9	Sequence 9, Appl1
32	59.4	1.3	4835	5	PCT-US96-02331-9	Sequence 9, Appl1
33	58.8	1.3	533	6	5482709-5	Patent No. 5482709
34	58.4	1.3	2201	4	US-09-330-970-2	Sequence 4, Appl1
35	58.4	1.3	3336	4	US-09-330-970-4	Sequence 4, Appl1
36	57.8	1.3	3254	2	US-08-965-903B-1	Sequence 1, Appl1
37	56.6	1.3	1332	2	US-08-481-814A-3	Sequence 3, Appl1
38	56.6	1.3	1489	3	US-08-836-582-1	Sequence 1, Appl1
39	56.4	1.3	3771	1	US-08-185-432-3	Sequence 3, Appl1
40	56.4	1.3	3771	1	US-08-185-432-23	Sequence 23, Appl1
41	56.4	1.3	5063	1	US-08-185-432-1	Sequence 1, Appl1
42	56.2	1.2	3819	1	US-07-686-322A-3	Sequence 3, Appl1
43	56.2	1.2	3819	1	US-08-002-999-3	Sequence 3, Appl1
44	56.2	1.2	7498	2	US-08-816-693A-1	Sequence 1, Appl1
45	56.2	1.2	7498	3	US-08-885-291-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-09-253-691-3
; Sequence 3, Application US/09253691
; Patent No. 6124100
; GENERAL INFORMATION:
; APPLICANT: Dong Kyu JIN
; TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
; TITLE OF INVENTION: Using Trinucleotide Repeats Sequence
; FILE REFERENCE: 1942/36
; CURRENT APPLICATION NUMBER: US/09/253, 691
; EARLIER FILING DATE: 1999-02-22
; EARLIER APPLICATION NUMBER: KR 98-6, 278
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: WordPerfect 6.1/Windows
; SEQ ID NO 3
; LENGTH: 397
; TYPE: DNA
; ORGANISM: human
US-09-253-691-3

Query Match      1.7%; Score 74.6; DB 4; Length 397;
Best Local Similarity 71.5%; Pred. No. 1.9e-11;
Matches 98; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 3756 AACGCACAGACAGCTGCTAGTCATCCTCCGACACAGAGGTCGCTATGATATGTC 3815
Db 108 aatgttcagcagcagcagcaaaagcagcagcagcagcagcagcagcagcagcagc 167

QY 3816 AGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGC 3875
Db 168 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 227

QY 3876 AACACAGCAACCCAG 3892
Db 228 agcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 244

RESULT 2
US-08-469-802B-4
; Sequence 4, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Rannum, Laura P.W.
; APPLICANT: Chung, Ming-Yi.
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5741645
```

```

? TITLE OF INVENTION: Type 1 and Method for Diagnosis
? NUMBER OF SEQUENCES: 47
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.
? STREET: 119 No. 5741645th Fourth Street, Suite 203
? CITY: Minneapolis
? STATE: MN
? COUNTRY: USA
? ZIP: 55401
? COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueeling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110,00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 168 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-4

Query Match          1.6% Score 73.2; DB 1; Length 168;
Best Local Similarity   69.7%; Pred.No.2,8e-11;
Matches    99; Conservative    0; Mismatches    43; Indels      0; Gaps     0;

QY  3768 AGCTGCTAAGTCAATCAGTCCTCCGACACAAGAGGGTGCTATGTATGCAGCAGCACG 3827
      |||         |  ||  |  ||  ||  |  |  |  |  |  |||||
Db   17  AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 76

QY  3828 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAAACAGCAACAGCAAGCAGCAGCAAA 3887
      |||||||
Db   77  AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 136
      |||||||

QY  3888 CCCAGGCCCTTCAGCCCACCCTCC 3909
      |||       |||||   |||
Db   137 AGCAGCACCTTCAGCAGGGCTCC 158

RESULT      3
US-08-267-803B-4
Sequence 4, Application US/08267803B
Patent No. 5834183
GENERAL INFORMATION:
APPLICANT: Ott, Harry T.
APPLICANT: Rannum, Laura P.W.
APPLICANT: Chung, Ming-yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene sequence for spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:

```

```

MEDUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentlin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267, 803B
FILING DATE: 28-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MCCORMACK, MYRA H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 110.00030120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 168 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-267-803B-4

Query Match 1.6%; Score 73.2; DB 2; Length 168;
Best Local Similarity 69.7%; Pred. No. 2.8e-11;
Matches 99; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3768 AGCTGCTAAGTCATCTTCCTCCGACACAGAGGGTGGCTATGATGTGCAGCAGCAGCAAC 3827
Db 17 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 76
QY 3828 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3887
Db 77 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 136
QY 3888 CCCAGGCTTCAGCCCACTCC 3909
Db 137 AGCAGCAGCCTCAGCAGGCTCC 158

RESULT 4
US-08-469-802B-5
Sequence 5, Application US/08469802B
Patent No. 5741645
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Rannum, Laura P.W.
APPLICANT: Chung, Ming-yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5741645
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Muelting, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 No. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentlin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469, 802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muelting, Ann M.

```

REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-5

Query Match 1.6%; Score 73.2; DB 1; Length 171;
Best Local Similarity 69.7%; Pred. No. 2.9e-11;
Matches 99; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3768 AGCTGCTAGTCACTCTCCGACACAGAGGGTGCTATGATGACAGCAGCAGCAAC 3827
DB 20 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 79
QY 3828 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3887
DB 80 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 139
QY 3888 CCCAGGCTTCAGCCACCTCC 3909
DB 140 ACAGCAGCCTCAGCAGGCTCC 161

RESULT 5
US-08-267-803B-5
Sequence 5, Application US/08267803B
Patent No. 5834183
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Rannum, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Muelting, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267.803B
FILING DATE: 28-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 110.00030120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-267-803B-5

Query Match 1.6%; Score 73.2; DB 2; Length 171;
Best Local Similarity 69.7%; Pred. No. 2.9e-11;
Matches 99; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3768 AGCTGCTAGTCACTCTCCGACACAGAGGGTGCTATGATGACAGCAGCAGCAAC 3827
DB 20 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 79
QY 3828 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3887
DB 80 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 139
QY 3888 CCCAGGCTTCAGCCACCTCC 3909
DB 140 ACAGCAGCCTCAGCAGGCTCC 161

RESULT 6
US-08-469-802B-2
Sequence 2, Application US/08469802B
Patent No. 5741645
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Rannum, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5741645
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Muelting, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 No. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469.802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muelting, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 195 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-2

Query Match 1.6%; Score 73.2; DB 1; Length 195;
Best Local Similarity 69.7%; Pred. No. 3.1e-11;
Matches 99; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3768 AGCTGCTAGTCACTCTCCGACACAGAGGGTGCTATGATGATGACAGCAGCAGCAAC 3827

[illegible]

```

RESULT      8
US-08-469-802B-3
; Sequence 3, Application US/08469802B
; Patent No. 5741645
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Rannum, Laura P.W.
; APPLICANT: Zonghl, Ming-yi
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5741645
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mueelling, Raasch, Gebhardt & Schwappach, P.A.
; STREET: 119 No. 5741645th Fourth Street, Suite 203
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469, 802B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueelling, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110,00030101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; FAX: 612-305-1225
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-469-802B-3

Query Match          1.6% Score 73.2; DB 1; Length 234;
Best Local Similarity 69.7%; Pred. No. 3.5e-11;
Matches   99; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY  3768 AGCTGCTAAGTCATCAGTCCGACAAAGAGGGTGCTATGCATGCGAGCAGCAAGC 3827
    ||| || | || | | |||| | | | | | | | | | | | | | | | | | | | |
DB  83 AGCAGCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 142
    ||| || | || | | |||| | | | | | | | | | | | | | | | | | | | |

QY  3828 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAACAGCAGCAGCAACAAGC 3887
    ||| || | || | | |||| | | | | | | | | | | | | | | | | | | | |
DB  143 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 202
    ||| || | || | | |||| | | | | | | | | | | | | | | | | | | | |

QY  3888 CCCAGGCTTAGGCCACCCTCC 3909
    ||| | |||| | ||||
DB  203 AGCAGCAGCCTCAGCAGGCGCTCC 224
    ||| | |||| | ||||

RESULT      9
US-08-267-803B-3
; Sequence 3, Application US/08267803B
; Patent No. 5834183
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Rannum, Laura P.W.
; APPLICANT: Chung, Ming-yi
; :

```

```

1 COUNTRY: USA
2 ZIP: 55401
3
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patent Release #1.0, Version #1.25
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/08/469,802B
11 FILING DATE: 06-JUN-1995
12 CLASSIFICATION: 435
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Mueeling, Ann M.
15 REGISTRATION NUMBER: 33,977
16 REFERENCE/DOCKET NUMBER: 110,00030101
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: 612-305-1217
19 TELEFAX: 612-305-1225
20 INFORMATION FOR SEQ ID NO: 6:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 154 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: single
25 TOPOLOGY: linear
26 MOLECULE TYPE: DNA
27 US-08-469-802B-6
28
29 Query Match 1.5%; Score 68.4; DB 1; Length 154;
30 Best Local Similarity 71.4%; Pred. No. 6,1e-10;
31 Matches 90; Conservative 0; Mismatches 36; Indels 0; Gaps 0
32
33 QY 3767 GAGCTGCTAAGTCATCACTTCGACACAGAGAGGGTGGCTATATGTCAGCAGCAGCA 3826
34 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 Db 2 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 61
36
37 QY 3827 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 3886
38 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 Db 62 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 121
40
41 QY 3887 ACCCG 3892
42 |||
43 Db 122 CAGCAG 127
44
45 RESULT 11
46 US-08-267-803B-6
47 Sequence 6, Application US/08267803B
48 Patent No. 5834183
49 GENERAL INFORMATION:
50 APPLICANT: Orr, Harry T.
51 APPLICANT: Rannum, Laura P.W.
52 APPLICANT: Chung, Ming-yi
53 APPLICANT: Zoghbi, Huda Y.
54 TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
55 Patent No. 5834183
56 TITLE OF INVENTION: Type 1 and Method for Diagnosis
57 NUMBER OF SEQUENCES: 85
58 CORRESPONDENCE ADDRESSES:
59 ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.
60 STREET: P.O. Box 581415
61 CITY: Minneapolis
62 STATE: MN
63 COUNTRY: USA
64 ZIP: 55458-1415
65 COMPUTER READABLE FORM:
66 MEDIUM TYPE: Floppy disk
67 COMPUTER: IBM PC compatible
68 OPERATING SYSTEM: PC-DOS/MS-DOS
69 SOFTWARE: Patent Release #1.0, Version #1.25
70 CURRENT APPLICATION DATA:
71 APPLICATION NUMBER: US/08/267,803B
72 FILING DATE: 28-JUN-1994

```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McCormack, Myra H.
: REGISTRATION NUMBER: 36,602
: REFERENCE/DOCKET NUMBER: 110,00030120
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-305-1217
: TELEFAX: 612-305-1228
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 154 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-267-803B-6

Query Match 1.5%; Score 68.4; DB 2; Length 154;
Best Local Similarity 71.4%; Pred. No. 6,1e-10;
Matches 90; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 3767 GAGCGGCAAGTCACTTCCGACACAGAGGGTGGCTATGATGACAGACAGCA 3826
Db 2 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 61
QY 3827 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAACAGCAGCA 3886
Db 62 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 121
QY 3887 ACCCAG 3892
Db 122 CAGCAG 127

RESULT 12
US-08-469-802B-7
: Sequence 7, Application US/08469802B
: Patent No. 5741645
: GENERAL INFORMATION:
: APPLICANT: Orr, Harry T.
: APPLICANT: Rannu, Laura P.W.
: APPLICANT: Chung, Ming-yi
: APPLICANT: Zoghbi, Huda Y.
: TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
: Patent No. 5741645
: TITLE OF INVENTION: Type 1 and Method for Diagnosis
: NUMBER OF SEQUENCES: 47
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
: STREET: 119 No. 5741645th Fourth Street, Suite 203
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/469,802B
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mueling, Ann M.
: REGISTRATION NUMBER: 33,977
: REFERENCE/DOCKET NUMBER: 110,00030101
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-305-1217
: TELEFAX: 612-305-1225
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 506 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-469-802B-7

Query Match          1.5%: Score 68.4; DB 1: Length 506;
Best Local Similarity 66.0%; Fred. No. 1.2e-09;
Matches 99: Conservative 0; Mismatches 51; Indels 0; Gaps 0;

OY 3760 CAGCAGAGACGCTCAACATCATCTCCAGACAAGAGGGGTGATTGATGCAGCA 3819
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 105 CAACATGGCGACTCTGAGGCCAGAGCCCGGGACACAAGGCTGAGCAGCAGCAGCAGCA 164
      ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||

OY 3820 GCAGCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGAACAAGCAGCAGCA 3879
      ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
DB 165 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 224
      ||||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||

OY 3880 GCAGCAAACCAGCGCCTTCAGCCCCACCTTC 3909
      ||||| ||| | |||| | |||| | |||| |
DB 225 GCAGCACCGCAGCAGCAGCAGCAGCGGCTCC 254
      ||||| ||| | |||| | |||| | |||| |

RESULT 13
US-08-267-803B-7
: Sequence 7, Application US/08267803B
: Patent No. 5834183
: GENERAL INFORMATION:
: APPLICANT: Orr, Harry T.
: APPLICANT: Rannum, Laura P.W.
: APPLICANT: Chung, Ming-yi
: APPLICANT: Zoghbi, Huda Y.
: TITLE OF INVENTION: Gene sequence for Spinocerebellar Ataxia
: Patent No. 5834183
: TITLE OF INVENTION: Type 1 and Method for Diagnosis
: NUMBER OF SEQUENCES: 85
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Mueller, Raasch, Gebhardt & Schwappach, P.A.
: STREET: P.O. Box 581415
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55458-1415
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/267,803B
: FILING DATE: 28-JUN-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: McCormack, Myra H.
: REGISTRATION NUMBER: 36,602
: REFERENCE/DOCKET NUMBER: 110,00030120
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-305-1217
: TELEFAX: 612-305-1228
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 506 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-267-803B-7

Query Match          1.5%: Score 68.4; DB 2: Length 506;
Best Local Similarity 66.0%; Fred. No. 1.2e-09;
Matches 99: Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 29, 2001, 17:49:17 ; Search time 147.55 Seconds
(without alignments)
8690.439 Million cell updates/sec

Title: US-09-041-994-1

Sequence: 1 GCTGATGCTGCTGACTCAGAG.....CATTTGACGAGATTTCTAG. 4496

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 195820 seqs, 142601680 residues

Total number of hits satisfying chosen parameters: 391640

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*

1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4434.2	98.6	4789	US-09-196-296B-1	Sequence 1, Appl
2	145	3.2	3361	US-09-196-296B-3	Sequence 3, Appl
3	68.4	1.5	521	US-09-815-343-422	Sequence 422, Ap
4	62.8	1.4	588	US-09-801-833-5367	Sequence 5367, Ap
5	60.6	1.3	1091282	US-09-335-032-12218	Sequence 12218, A
6	57.8	1.3	407	US-09-724-866A-24873	Sequence 24873, A
7	57.6	1.3	924430	US-09-335-032-12216	Sequence 12216, A
8	53.8	1.2	619	US-09-724-866A-12926	Sequence 12926, A
9	53.6	1.2	2685	PCT-US01-08117-154	Sequence 154, Ap
10	53.2	1.2	45333	US-60-248-505-604	Sequence 604, Ap
11	53	1.2	161	US-09-724-866A-8024	Sequence 8024, Ap
12	53	1.2	924430	US-09-335-032-12216	Sequence 12216, A
13	52.4	1.2	237	US-08-276-163D-12927	Sequence 12927, A
14	52	1.2	574860	US-09-335-032-12208	Sequence 12208, A
15	51.8	1.2	666448	US-09-335-032-12214	Sequence 12214, A
16	51.6	1.1	5630	US-09-709-126-1	Sequence 1, Appl
17	51.6	1.1	439885	US-09-335-032-12212	Sequence 12212, A
18	51.4	1.1	784328	US-09-335-032-12217	Sequence 12217, A
19	51.2	1.1	287	US-09-724-866A-9061	Sequence 9061, Ap
20	51.2	1.1	340	US-09-724-866A-10002	Sequence 10002, A
21	50.2	1.1	562638	US-09-335-032-12211	Sequence 12211, A
22	50	1.1	948061	US-09-335-032-12219	Sequence 12219, A
23	50	1.1	1531974	US-09-335-032-12207	Sequence 12207, A
24	49.4	1.1	3629	US-09-357-273A-1	Sequence 1, Appl
25	48.8	1.1	3046	US-09-783-514-1823	Sequence 1823, Ap
26	48.8	1.1	3046	US-09-783-514-1823	Sequence 1915, Ap
27	48.8	1.1	813137	US-09-335-032-12205	Sequence 12205, A

28	48.6	1.1	293	US-09-724-866A-15386	Sequence 15386, A
29	48.6	1.1	324	US-09-724-866A-19400	Sequence 19400, A
30	48.2	1.1	1091282	US-09-335-032-12218	Sequence 12218, A
31	47.6	1.1	391	US-09-487-566A-5503	Sequence 5503, Ap
32	47.2	1.0	183	US-09-724-866A-12068	Sequence 12068, A
33	46.6	1.0	270148	US-09-335-032-12209	Sequence 12209, A
34	46.4	1.0	471	US-09-801-833-1504	Sequence 1504, Ap
35	46.2	1.0	402	US-09-724-866A-4428	Sequence 4428, Ap
36	46.2	1.0	948061	US-09-335-032-12219	Sequence 12219, A
37	46	1.0	813137	US-09-335-032-12205	Sequence 12205, A
38	45.8	1.0	229	US-09-540-212A-32189	Sequence 32189, A
39	45.8	1.0	315341	US-09-335-032-12206	Sequence 12206, A
40	45.4	1.0	313	PCT-US01-10661-5	Sequence 4358, Ap
41	45.4	1.0	3132	US-09-533-077-320	Sequence 320, Ap
42	45.4	1.0	3132	US-09-739-445-3547	Sequence 3547, Ap
43	45.2	1.0	378	US-09-775-861A-54	Sequence 54, Appl
44	45.2	1.0	448	US-09-739-449-188	Sequence 188, Ap
45	45.2	1.0	9152	US-09-739-449-188	Sequence 188, Ap

ALIGNMENTS

RESULT	1
US-09-196-296B-1	
Sequence 1, Application US/09196296B	
GENERAL INFORMATION:	
APPLICANT: Suen, Chen-Shian	
APPLICANT: Fryll, Donald E.	
APPLICANT: Lytle, Richard C.	
TITLE OF INVENTION: Cloning and Expression of a Nuclear	
FILE REFERENCE: 0630/01376	
CURRENT APPLICATION NUMBER: US/09/196, 296B	
CURRENT FILING DATE: 1998-11-19	
NUMBER OF SEQ ID NOS: 3	
SOFTWARE: FastSeq for Windows Version 3.0	
SEQ ID NO 1	
LENGTH: 4789	
TYPE: DNA	
ORGANISM: Homo Sapien	
US-09-196-296B-1	
Query Match	98.6%; Score 4434.2; DB 5; Length 4789;
Best Local Similarity	99.3%; Pred. No. 0;
Matches 4483; Conservative	0; Mismatches 8; Indels 24; Gaps 2;
QY	1 GCTGATGCTGCTGACTCAGAGACCAATTAATACTGCTGAACATCCCTTGACGCTT
DB	100 gctgatgtgtgagactcagagacccaataaataactgttgaacatcccttgaactggt
QY	61 AGCCATTTGCTGATGCTATATTTCAAGATGATGATTAGAGAAACTTGATCCTGCTGC
DB	160 agccagtggtgtgtatattcaagatgattagagaaacttgatcctgctgc
QY	121 CAGTATTCAGAAAGCAAAATTCATGCTGATCTCCAGCAAGCTTACTCTGACG
DB	220 cagtgatccagaaacgcaaatcgcattgattccaggaagcttccctgcag
QY	181 TGTGTAAGAAAGCAAGCGAGGAGCAAGTAATTAATGAAATTTGGCTGAGCTGAT
DB	280 tgttgaagaaacgagagcaggaagcaagaaatlaattgaagaatcgctgaactgat
QY	241 ATCTGCAATCTTAGATGATATTGACATTTCAATGTCMAACGATTAATGTCGATTTT
DB	340 atctgcacatcttagatattgacatttcaatgcaacacgataatgctgattt
QY	301 AAGGAAACAGTAAGACAGATACGTCATTAATAAGCAAGCAAGCAAGCAATTTTCATGA
DB	400 aaggaacagtaagacagatagcgtcaataaagcaagcaagcaagcaatatttccaatga
QY	361 TGATGATGTTCAAAAGGCGCATGCTTCTTACAGGCGAGGAGGCTTATGATTAACATC

|||||
Db 460 tgaatgtctcaaaaaagccgcatgatatctctcaaggcgaggtatattgataaagacc 519
OY 421 ctttagaccgcttttttacctgagcatttgatggtttcctattttgtgtgaatcgagagcc 480
Db 520 cttagagccgcttttaactcagcatctgagtcttcctactctggtgagatcagagcg 579
OY 481 AAAcATTGATTGTATCGAAAAATGTCAcAAATACCTCGCAATATAGCAAGAGACCTT 540
Db 580 aaacattgattctgtatccgaaaaatgctcacaaataccgtcaataataagaaagaaact 639
OY 541 GGTAAACAAGTGTATACATATCTTACATGAGAAGACAGAAAAGATTTTCTTAAGA 600
Db 640 ggttaacacaagtgcttaacaatacttaacaagaagacgaagaagattctcttaagaa 699
OY 601 TTTACCAAAATCTACAGTTTAAATGAGTTTCGTGCAAAATGAGCCCAAAACAAAAA 660
Db 700 tttaacaaaactacagtttaagtgaggttctctggaacaaatgagacccaagacaaaaa 759
OY 661 CCATACATTTTATTTGCCGATGTGATGAAAAACACCACTGATATTCTTGAAAGACATAA 720
Db 760 ccaatacattaaattgcgtatgttgatgaaaaacacacacatgaatcttgaaagacataa 819
OY 721 GSCAGTCCCTGAATGCGCCAGAGATATGAACAATGCAGTGTGCTTGCTCAGCC 780
Db 820 cgcaggtctctgaaatgycgcagagataatgaacaaatgcagtgcttgccctgctcaagcc 879
OY 781 ACGAGCTATGATGAGGAAGGGAAGATTGCAATCTTGATGATCTGTGTGGACGCGC 840
Db 880 acgagcattgagtgaaggagggagatttgcaatcttgatgactctggtgagcgccg 939
OY 841 CATTACTACAGAGAAAGACATTTCCATCAAAACCTGAGACTTTATTACAGACATGA 900
Db 940 catctcctcaggaagaagaacattccacaaacctgagagccttattacagacatba 999
OY 901 TCTTTCAGAAAGTGTGTCAATATAGATATCAAAATTCACGAGATCTCCATGAGGCTTG 960
Db 1000 tcttccaggaagaagtgctgaataatagatacaaatccacgagatctccatgagagccttg 1059
OY 961 CTTTGAAGATATTAATCCGAAGGTGTAATTCAGAGATTTTAACTCTTAATGATGGCAGTC 1020
Db 1060 ccttgaaagataaacacgaaggctgatacagatcttcttaagctcaaatgagtgagcagtc 1119
OY 1021 ATGGTCCCAAGAAAGCTACTATCAGAAGGCTTATCTTAATGGCCATGCAAGAACCCCAT 1080
Db 1120 atggtccccaagaaacgtccactacaaagaagcttacttaagtgccatgcgaagaccccgat 1179
OY 1081 ATATGATTCTCTGTGGCTGATGGAATATAGTGACTGCACAGACAAAAAGAACTCTT 1140
Db 1180 atatcgattctcgctggtcgtgatactatagtgactgcacagacaaaaagcaaatctct 1239
OY 1141 CCGAAATCTGTACAAATGATGCAGATGGCTTTGTCTCAACCCACTTCCTTCAAGAGAC 1200
Db 1240 ccgaatctcgttaacaaatgatacgaatgcttgctcaacccactctctccacagaga 1299
OY 1201 ACAGATGGATATAGACCAAAACCAAAATCTGTGGACAAGGATATAGACCACTATGGC 1260
Db 1300 acaagaatggaatacagaaccaaacaccaaatccgtctggaacaaaggaatlaagccactatggtc 1359
OY 1261 TGGATGCAACAGTTGCTGATAGCGGCGATGAGTATGTCTGCCAAACCAAGGTTACAGATGCC 1320
Db 1360 tggatgcaaacagctcgttagagcgcatgagtatgtcgccaacaaagcttacaagatgctc 1419
OY 1321 GAGCAGCAGGGCCTTATGGCTTGGCAGACCCCTTAGACACACAGGCGAGATGATGAGCTAG 1380
Db 1420 gagcagcgagcgctatgcttgcgagacccctagcaccaagcgagatgagtgagagctag 1479
OY 1381 GATNGGGGTTCCAGTAACATAGCTTTCATTTGACCCCTGGCCAGGACATGCATACCAATC 1440
Db 1480 gtaatgggggtctccagtaaaatagctcatctgacccctggcgcaagcatgcaatccacatc 1539
OY 1441 TTCTTACCAAGAACCAACTATGGGCTCAACATGATGAGCCCCCAACATGGAGTCTTGG 1500
|||||

Db 1540 ttccctacccaagaacaactatggygtcacaatgagtgagccccccacatggyagtccg 1599
OY 1501 TCTTCCCCCAAAACAGCAGACAATATCATGATTTTCTCTCGTAATCTGGGAGTCAAGAAAT 1560
Db 1600 tcttgcccaaacccagcagaatatacatgattctctctctgtaatctggtgagtgcaaatg 1659
OY 1561 AGCCTCACATCAGTTTCTCTCTGTTCGAGGTGTGCATCTCCCATGCGCATCTTCTGGCAA 1620
Db 1660 agccctcacatcagttctctccctgtgcaggtgtgcactctccacatggaatctctcgcga 1719
OY 1621 TACTGGAAACACAGCTTTTCCAGCAGCTCTCTAGTGGCCCGCAAGCATAGGAAG 1680
Db 1720 tactggyaacccaagcttctccagagctctctcagtgctccgtgcaagccaatcagtgaaag 1779
OY 1681 TGTGGGACTTCCCTTTTATCTACTCTGTGCATCACCAGGCCCAAAATTGATTAATCTTCC 1740
Db 1780 tgtggyagcttcccttatactctctgtcatcaccaagcccccaatctgataactctcc 1839
OY 1741 CAATATGATATTTACCCAACCAAGTAAGCTAAGCAATCAGATTTCCAAAGTCTCTGGG 1800
Db 1840 caataatgaatatacccaaccaaagtaagcaatcaggaatccaaagatcccttg 1899
OY 1801 CTTTATTTGCGACCAAAATCCAGTGAGAGTTCAATGTGTCAGTCAAAATPACGAGATCA 1860
Db 1900 ctttatactgagccaataatccagtggaagatccaatgtgtcagttcaaatagcagaatca 1959
OY 1861 CCTCAGTCAAAAGAAAGTAAGAGAGCAGTGTGAGGGGCGCAGAAATCAAAAGGGTCC 1920
Db 1960 cctcattgacaagaagaagtaaggaagcaggtgtgagggggagagaatccaagggtccc 2019
OY 1921 TTTGGAAAGCAAAAGTCAATAAAAATTTACTGCAATTACTTACCTGTTCTTCTGATGACC 1980
Db 2020 ttggaagaacaaaggtcataaaaaatctactcagttacttaactctgtctctcgtgacag 2079
OY 1981 GGGTATTCCTCTTTGACCAACTCCCCCTAGATTTCAAGTTTAAAGATCTTCTGTAG 2040
Db 2080 ggttacttctctctgacccaactccccctgatactcaagtgtgaaagaatctctcgttag 2139
OY 2041 TGTCAACAGCCCTCTGTGAGTCTCTCTCTATCTGTGAAGAGTATCTTACATCTCAA 2100
Db 2140 tgtcaacagccctctgagttctctctctactacatctgaaagatcatctctcaatccaa 2199
OY 2101 TATGATGGGTACAGTTTACAGAGACAGCGGATTTTGCAAGTGTGCTCAGAAATGG 2160
Db 2200 tatgcatggttactcgttatacaagaagaacccggaatcttgcaagatctgctgcagaatg 2259
OY 2161 GAATTCACACAGCTGAGTGAAGCCCAAGATTACTGCACAAAGCCACTGGGAAAGACACAGCAG 2220
Db 2260 gaattcacacagctgagtgatgccaagatactactgcagaagccactggyaaagacacacagcag 2319
OY 2221 TATTAATCTTGTGGGACGGAATATGTTGTCAAGCAGAGCAGCTTAAGTCTTAAGAAGA 2280
Db 2320 tataactctctgttggygaagyaatctgttcaagaagagcagctaaagcttctaagaagaa 2379
OY 2281 GGAGATATATGACATTTTACATGATACCTGTGACAGGGATGTACTAGATGACACTCTC 2340
Db 2380 ggaagaataatgacatctcttaagatactctgcgcagaggaatgactctctatgagcactctc 2439
OY 2341 TAAAGAACTACAGCCCAAGTGAAGGAGTGACATATAAATGAGTCACTGACACAGCTC 2400
Db 2440 taagaacatcagccccaagtggaagagtggaataataaagtatcagtgacagatc 2499
OY 2401 CACCAATCTTCACTCAAGTCAAGAGAAAGACCCCTTAATAATTAAAGACAGACAAAGTGAAGA 2460
Db 2500 caccattctctagctcaagtcgaagagaagaaccttaaaatlaagacagagaacaagtgaaga 2559
OY 2461 GGGATCTGGAGCTTGGATTAATCTAGATGCTATCTTGTGTGATCTGACTAGTTGACTT 2520
Db 2560 gggatctggagactctggaataactagatgctacttctgtgtgactcgtgactgtctcgaact 2619
OY 2521 TTACATTAATTCATATCTCTCAATGATGATGATCTGGGAGATTAAGCAACAGGTCTTCA 2580
Db 2620 ttacataaatctcatalctctcaaatctgtagtcaatctggygaactaaagcgaacaggtgttcca 2679


```

: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 3
: LENGTH: 3361
: TYPE: DNA
: ORGANISM: Murine
: OS-09-196-296B-3

```

Query Match	3.28;	Score 145;	DB 5;	Length 3361;
Best Local Similarity	84.5%;	Pred. No. 1.6e-27;		
Matches 163;	Conservative	0;	Mismatches 30;	Indels 0;
			Gaps	0;

Qy	427	ACCGCTTTTACTTCAAGGCATTGGANGTTTCTATTTTGTGTGTAATCGAAGGCAAACT	486
Db	2008	acacattccactgttaggcacgtgatgtttccctgtctgtggtgaatcgagatgaaacat	2067

Qy 487 TGTATTTCATCAGAAATGTCACACATACCTGCAATATTAAGCAGAGGACCTGGTTAA 546
||||| || ||||||||||||||| || |||| |||||||||||||
Db 2068 tgrattcgtgcagaaatgtcacacagtatctgcagtacaagcagcagcgacctggtta 2127

Oy	547	CACAGTGTTTACAATATCTTCATCATGAAGAACAACGAAAGGATTCTTTAAGCAATTTACC	606
Db	2128	cacaaagtgtctaacagcatcctatcaitgacaagaaccgaagaagtatttcctaaceacttacc	2187

Qy	607	AAAACTCTACAGTT	619
Db	2188	aaatccacagtc	2200

RESULT 3
US-09-815-343-422/C

```

; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: Xu, Jiangchun

```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURREN ADDITION NUMBER: 05/00/015 243

```

```
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 422
```

```

; LENGTH: 521
; TYPE: DNA
; ORGANISM: Homo sapiens
FEATURE:

```

```

; NAME/KEY: misc_feature
; LOCATION: (1)...(521)
; OTHER INFORMATION: n = A,T,C or G
MS-09-815-343-A23

```

Query Match	1.58;	Score 68.4;	DB 5;	Length 521.
Best Local Similarity	73.79;	Dred NO. 4	10-09;	

QY	3812	ATTCGACGCAGCAGCAAGCAGCAGCAGCAGCAGCAGCAAGCAAGCAAGCAAGCAG	3871
Matches	87; Conservative	0; Mismatches 31; Indels	0; Gaps

D_b 151 AAGCAGCAGCAGCAGCAGCAGCAGCACCTGCGGACAGCAGCAACAGCAGCAACAACAG 92
Qy 3872 CAGCAACAGCAGCAAAACCCAGGGCTTTCAGGCCAACCCTTAATGTGACTGCCTTCCCCCA 3929

Db 91 CAGCAACACAGCAGCAGCAGCAGCAGTAAAGGGCATACATTTCCTGCTTACCA 34

US-09-801-833-5367
; Sequence 5367, Application US/09801833
; GENERAL INFORMATION:
; APPLICANT: Glucksmann M Alexandra

10 TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A

```

; TITLE OF INVENTION:  HUMAN BRAIN LIBRARY
;
; FILE REFERENCE:  1600.1037-005
;
; CURRENT APPLICATION NUMBER:  US/09/801,833

```

; CONSENT FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 09/371,168
 ; PRIOR FILING DATE: 1999-08-10

PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/103,145

```

; NUMBER OF SEQ ID NOS: 8285
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 550
; LENGTH: 588
; TYPE: DNA

```

```

; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (1) : (500)
; OTHER INFORMATION: n = A, T, C or G
US-09-801-833-5367

```

Query Match	1.4%;	Score 62.8;	DB 5;	length 588;
Best Local Similarity	77.6%;	Pred. No. 1.1e-06;		
Matches	76;	Conservative	0;	Mismatches 22;
			Indels	0;
			Gaps	0;

Qy 3804 CTTATGATGATGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAACGCAAC 3863
 | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 87 ccattgtcgtgaagcccaagcagcagcagcagcagcagcagaacaagcagcagcaaac 146

QY	3864	AGCAACAGCAGCAACAGCAGCAAAACCCAGGCTTTCAGC	3901
Db	147	agcagcgcagcagcagcagcagccgcgcgcgcgcgcgc	184

RESULT 5
US-09-335-032-12218

```

; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert

```

```

; TITLE OF INVENTION: Characterization of the yeast
;
; TITLE OF INVENTION: Transcriptome
;
; FILE REFERENCE: 01107.78572

```

; CURRENT FILING DATE: 1999-06-16
 ; PRIOR APPLICATION NUMBER: US 60/035,917
 ;

PRIOR APPLICATION NUMBER: US 09/012,033
PRIOR FILING DATE: 1998-01-22

```

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12218

```

LENGTH: 1031202
 TYPE: DNA
 ORGANISM: *Saccharomyces cerevisiae*

Query Match 1.3%; Score 60.6; DB 5; Length 1091282;

	Matches	84;	Conservative	0;	Mismatches	39;	Indels	0;	Gaps	0.
QY	3807	TGATGATGCAGCAGCACACAGCAGCAGCAGCAGCAGCAGCAACAGCAACAGC	3866							

Db 236948 tgaana tgcagcagcaacaacacagcagcaacaacacagcagc 237007
QY 3867 AACACGACGACGACGCAACCCAGGCTTCAGGCCACCTCCTAATGTGACTGCTCCC 3926

Db 237008 aacacaacagcaacgacatatatccctccctgcgacitcctgtygtgycataattcgy 237067

[illegible]

Search completed: April 29, 2001, 21:12:00
Job time: 12163 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 16:45:52 ; Search time 32.49 Seconds
(without alignments)
836.669 Million cell updates/sec

Title: US-09-041-994-2

Perfect score: 7383
Sequence: 1 MSGIGENLDPLASDSRRKRL.....MNNPMMSGMPGPDQKRC 1415

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCROS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	4.2	855	2	US-08-816-693A-2
2	307	4.2	855	3	US-08-885-291-2
3	286.5	3.9	846	3	US-08-885-291-55
4	286.5	3.9	846	3	US-09-107-847-2
5	276.5	3.7	2703	1	US-08-185-432-19
6	259	3.5	2414	1	US-08-227-536-2
7	259	3.5	2414	5	PCT-US95-04682-2
8	258	3.5	826	1	US-08-785-241-6
9	258	3.5	826	2	US-08-480-473B-2
10	258	3.5	826	2	US-08-915-213-2
11	258	3.5	826	4	US-09-148-547-2
12	258	3.5	826	4	PCT-US96-10251-2
13	254.5	3.4	816	2	US-08-785-310A-8
14	254.5	3.4	816	2	US-08-816-693A-53
15	254.5	3.4	816	3	US-08-885-291-53
16	254.5	3.4	2441	1	US-08-194-468-2
17	254.5	3.4	2441	1	US-08-961-739-2
18	252	3.4	848	1	US-08-045-806-4
19	252	3.4	848	1	US-08-366-051B-4
20	249	3.4	805	2	US-08-480-473B-4
21	249	3.4	805	5	US-08-915-213-4
22	249	3.4	805	5	PCT-US96-10251-4
23	239	3.2	810	1	US-08-785-241-7
24	237.5	3.2	870	1	US-08-785-241-4
25	233.5	3.2	824	3	US-08-885-291-52
26	230.5	3.1	824	2	US-08-816-693A-52
27	226.5	3.1	824	2	US-08-785-310A-7

ALIGNMENTS

28	222.5	3.0	1142	2	US-08-993-118-7	Sequence 7, Appl
29	222.5	3.0	1142	3	US-08-845-528C-7	Sequence 7, Appl
30	221	3.0	1618	1	US-07-853-913-4	Sequence 4, Appl
31	217.5	2.9	875	1	US-08-785-241-5	Sequence 5, Appl
32	217	2.9	1663	5	PCT-US93-07261-16	Sequence 16, Appl
33	212.5	2.9	3969	4	US-08-061-376-5	Sequence 5, Appl
34	204.5	2.8	805	1	US-08-045-806-2	Sequence 2, Appl
35	204.5	2.8	805	1	US-08-366-051B-2	Sequence 2, Appl
36	203.5	2.8	788	2	US-08-918-914-4	Sequence 4, Appl
37	202.5	2.7	1093	3	US-08-545-860D-55	Sequence 55, Appl
38	202.5	2.7	1093	5	PCT-US94-04496-55	Sequence 7, Appl
39	202	2.7	2842	1	US-07-741-940-7	Sequence 7, Appl
40	202	2.7	2842	1	US-08-289-548A-7	Sequence 7, Appl
41	202	2.7	2842	1	US-08-452-654-7	Sequence 2, Appl
42	202	2.7	2843	1	US-07-741-940-2	Sequence 2, Appl
43	202	2.7	2843	1	US-08-289-548A-2	Sequence 2, Appl
44	202	2.7	2843	1	US-08-452-654-2	Sequence 2, Appl
45	202	2.7	2843	1	US-08-452-655B-2	Sequence 2, Appl

RESULT 1
US-08-816-693A-2
Sequence 2, Application US/08816693A
Patent No. 5874241
GENERAL INFORMATION:
APPLICANT: Takahashi, Joseph S
APPLICANT: Turek, Fred W
APPLICANT: Pinto, Lawrence H
TITLE OF INVENTION: Clock Gene and Gene Product
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESS: Dressler, Rocky, Milnamow & Katz
STREET: Two Prudential Plaza, Suite 4700
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/816,693A
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5874241thrup, Thomas E
REGISTRATION NUMBER: 33,268
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5400
TELEFAX: 312-616-5460
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-816-693A-2
Query Match 4.2%; Score 307; DB 2; Length 855;
Best local Similarity 19.9%; Pred. No. 2,4e-14;
Matches 220; Conservative 149; Mismatches 373; Indels 364; Gaps 48;
QY 34 EKRRROESKYTEELAEILSANISDIDNENYKDKCAILKEIVYRQIRQIKEDGKTSND 93
DB 43 EKRRROOFNVLKELISMLPGNAR-----KKDKSTVLRKSIDFLRKHKKE--TTAQSDA 93
QY 94 DVGKAVSVSGGQGVIDKSLGPIILQALDGFLEFVVRANIVSVSENVYQYLYQKQEDLV 153

Db 94 SEIRQDMKPT---FLSNBEFTQMLEALDGFELAMTDSITIVSESVTSLEHLPSDLV 150
QY 154 NTSVYNILHEEDRKDFELNPKSTVNGVSWTNEPOROKSH--TFNCRYMLMKT--PHD----- 206
Db 151 DOSIFNFPBEGHSHVYKILSTHLESLSLTPREYLSKSNQLEFCCHMLRGITIDPREPSTY 210
QY 207 -----ILEDINASPE-----MRORYETWQCFR---LSQPRAMMEEGEDL 242
Db 211 EYVRFIGNFKSLTSTVSTHNGFEETIOGTRHPSYEDHVCFAVATVRLATPOFIKE----- 265
QY 243 QSCMICVARRITGTGERTPPSPNESTIRHDLGKVVNIDTNS-----LRSSMRP 291
Db 266 ---MC-----TVEEPNEFTSRHSLEWKFLFLDHRAPPIIGYLPREVIGTS--- 308
QY 292 GFEDIIRRCIORFFSLNDGQSWQKRHYQEAUYNLNGHAETPVYRFRSLADGTIVTAQTKSKL 351
Db 309 GYD-----YHVDDLENLAKCHEHLMQY--CKGKSCYYRFLTKQOQMIWQIQT----- 353
QY 352 FRNPVTNDRHGCVSTHFLQREONGYRPNPNVGGGIRPPMAGCNSSVGMGSMSPNOGLQW 411
Db 354 -----HYIITVH-----QWNS-----RPEFIVCHTIVVSYA-----EVRA 383
QY 412 PSSRAYGLAD--PSTTGMGSGARYGSS--NIASL-----TPRGQMSPSYQNN 457
Db 384 ERRRELGLIEESLPETPADKS--QDSGSDNRITVYSLKALELRFHSPSPASSRSRKS 441
QY 458 NYGLNMSPPHSGPLANOONIMISPRNRSKPRIASHQFSPVAGVHSPMASSGNTGNGHS 517
Db 442 HTAV---SDPSSTPYKIFTDTS--TPRQO-----HLPAHEKMTQRSS 479
QY 518 FSSSLSLALQAISEGVGSLSTLSSPGPKLDNSPNMNITOPSKVSNDSKSPLEGFYCDQ 577
Db 480 FSSQSIN-----SQSVGSLTQ-----PAMSOQANLPI--POGMSQOFOQAQG----- 521
QY 578 NPVSSMCQNSRDLHSKESKESSEVEGAENORGPLESKGGHKKLIQTLTCSDDRGHSSL 637
Db 522 -----AMQHLKQLEORTMIRANIHROQELRKIOEOLQ----- 557
QY 638 TNSPLDSSCKESSVSVTSPGVSSTSGVSSSTSMHGSLLQEKHRIHLKLLONGNSPAE 697
Db 558 -----VHGGQIQ-----MFLQOSN----- 571
QY 698 VAKITAQATGKDTSI--TSCDGNVYKQEOLSPPKKENALLRYLDDDDPSDALSKELQ 756
Db 572 -----PGLNFGSVQJLSSGNSNI--QQLTP-----VNMQ 597
QY 757 POVEGVDNKMQCTSTTIPSSQOEKDPKIKTETSEEGSGDLDNLDAIGDLTSSDFYNN 816
Db 598 GQV-----VPANQVQ-----SGHI----- 611
QY 817 ISSNSHLGTQOQVFOGTNSLGLKSSQVOSIRPPYNRAVSLDSPVSGSSPPVKNISAF 876
Db 612 --STGOHM-100QTLQSTNS-----TQOSQOSVMSGHSQOTSLPQSTPLTFLAYN----- 659
QY 877 PMLPQOPMLGPRKMDQOENGSSMGGRNKNRVITYTOTPSSCDMGLPMSKAGRMPPMNSN 936
Db 660 TMVISOAPAG-----SMVOIPSSM--PONTQOSATVTTFTQDROIRISQO---QQLVTK 708
QY 937 SMGRPGCDYNTSLPRPALGSI---PTLPLRSNIPGARPLYLOQOQOQMLQMRPEIDMG 992
Db 709 LVTPAVACGAVVNPSTMLMGQVYVATPTATQOQOQO--TLSTVTOOQOQOQOQOQPPQOQOQ 767
QY 993 MGAN-----PYGQAASQNLGSMWPDGMLSMEOVSHGTGNRPLLRNS---LDDLVPSPNL 1044
Db 768 QOSSQEQQLPSVQOPAQOQLOPQOQFLOTGSRHLHGNSPTQJLISAAVPLQOSTPSPSHH 827
QY 1045 EGQSDERALLDQJLHTLSNTDATGLE 1070
Db 828 QOHOPQOQOQQLPRHRTDSLTPSKVQ 853

US-08-885-291-2
; Sequence 2, Application US/08885291A
; Patent No. 6057125
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/08/885,291A
; CURRENT FILING DATE: 1997-06-30
; EARLIER APPLICATION NUMBER: 08/816,693
; EARLIER FILING DATE: 1997-03-13
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRF
; ORGANISM: Mus musculus
US-08-885-291-2

Query Match 4.2%; Score 307; DB 3; Length 855;
Best Local Similarity 19.9%; Pred. No. 2.4e-14;
Matches 220; Conservative 149; Mismatches 373; Indels 364; Gaps 48;

QY 34 EKRREQSKYIEELAEELISANLSDIDNFNVKPKCALIKETVROIROKEQKRTISND 93
Db 43 EKRREQSKYIEELAEELISANLSDIDNFNVKPKCALIKETVROIROKEQKRTISND 93
QY 94 DYQKADVSTGQGVADKSLRPLLIQALDGFVYVNRANIVFSENVTOYLOKQOBYL 153
Db 94 SEIRQDMKPT---FLSNBEFTQMLEALDGFELAMTDSITIVSESVTSLEHLPSDLV 150
QY 154 NTSVYNILHEEDRKDFELNPKSTVNGVSWTNEPOROKSH--TFNCRYMLMKT--PHD----- 206
Db 151 DOSIFNFPBEGHSHVYKILSTHLESLSLTPREYLSKSNQLEFCCHMLRGITIDPREPSTY 210
QY 207 -----ILEDINASPE-----MRORYETWQCFR---LSQPRAMMEEGEDL 242
Db 211 EYVRFIGNFKSLTSTVSTHNGFEETIOGTRHPSYEDHVCFAVATVRLATPOFIKE----- 265
QY 243 QSCMICVARRITGTGERTPPSPNESTIRHDLGKVVNIDTNS-----LRSSMRP 291
Db 266 ---MC-----TVEEPNEFTSRHSLEWKFLFLDHRAPPIIGYLPREVIGTS--- 308
QY 292 GFEDIIRRCIORFFSLNDGQSWQKRHYQEAUYNLNGHAETPVYRFRSLADGTIVTAQTKSKL 351
Db 309 GYD-----YHVDDLENLAKCHEHLMQY--CKGKSCYYRFLTKQOQMIWQIQT----- 353
QY 352 FRNPVTNDRHGCVSTHFLQREONGYRPNPNVGGGIRPPMAGCNSSVGMGSMSPNOGLQW 411
Db 354 -----HYIITVH-----QWNS-----RPEFIVCHTIVVSYA-----EVRA 383
QY 412 PSSRAYGLAD--PSTTGMGSGARYGSS--NIASL-----TPRGQMSPSYQNN 457
Db 384 ERRRELGLIEESLPETPADKS--QDSGSDNRITVYSLKALELRFHSPSPASSRSRKS 441
QY 458 NYGLNMSPPHSGPLANOONIMISPRNRSKPRIASHQFSPVAGVHSPMASSGNTGNGHS 517
Db 442 HTAV---SDPSSTPYKIFTDTS--TPRQO-----HLPAHEKMTQRSS 479
QY 518 FSSSLSLALQAISEGVGSLSTLSSPGPKLDNSPNMNITOPSKVSNDSKSPLEGFYCDQ 577
Db 480 FSSQSIN-----SQSVGSLTQ-----PAMSOQANLPI--POGMSQOFOQAQG----- 521
QY 578 NPVSSMCQNSRDLHSKESKESSEVEGAENORGPLESKGGHKKLIQTLTCSDDRGHSSL 637
Db 522 -----AMQHLKQLEORTMIRANIHROQELRKIOEOLQ----- 557
QY 638 TNSPLDSSCKESSVSVTSPGVSSTSGVSSSTSMHGSLLQEKHRIHLKLLONGNSPAE 697
Db 558 -----VHGGQIQ-----MFLQOSN----- 571

```

OY 698 VAKITAAATGADGTSSI--TSGCDGVMVQKQELSPKKKENNALLRYLLDRDPDALSKEIQ 756
Db 572 -----PGLNFSGSVSSGNSNYI--QQLTP-----VNNQ 597
OY 757 PVEGVDVNNKMSQTSSTYIPSSOEKDPKIKTERSEEGCDLMDLIDLGLTSSDFPNYS 816
Db 598 GQV-----VPAHQVQ-----SGHI-----611
OY 817 ISSMSGLGTQKQVFGCTNSLGLKSSQSVOSIRPPYNRAVSLDSPVSGSSPPVKNISAF 876
Db 612 --STGQHM-100QTLQSTST-----TQSSQSVMSGSHSQSTSLPSTQTPSLTAPLVN--- 659
OY 877 PMLCKPMPLGSPRMMDMSQENYSSMSGCPRRNTVYQTSSGDMGLPNSKAGRMEMNSN 936
Db 660 TMTVTSQPAAG-----SMVOIPBSM-PQNSTQSAVTYTFDQRLRFSGQ---QQLVTK 708
OY 937 SMGRPGGDYNTSLPRPALGSI-----PTLRGRNSIIPGAPRYLQOQOQMLQMRPGEPIMG 992
Db 709 LVTAPVACGAMVPSMTMLMQVYVATFTATQOQQAQ-FLSYVQOQOQOQOQOQPPQOQOQO 767
OY 993 MGAN-----PYGQAASNOGLSMPDGLMSKEQVSHGCTQNRPLRNS---LDDLVGPPSNL 1044
Db 768 QOSSOEQOLLEPSVOQPAQAOGLQPPQOFLQTSRLHLGNPSQTLLSAPFLQSTFPPSHH 827
OY 1045 EGSGDERALLDQQLHTLLSNIDATGLE 1070
Db 828 QQHQPQOQOQQLPRHRTJSTLDPKVKO 853

```

```

? RESULT 3
? US-08-885-291-55
? Sequence 55, Application US/08885291A
? Patent No. 6057125
? GENERAL INFORMATION:
? APPLICANT: Takahashi, Joseph S.
? APPLICANT: Turek, Fred W.
? APPLICANT: Pinto, Lawrence H.
? TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
? FILE REFERENCE: 0290-5
? CURRENT APPLICATION NUMBER: US/08/885,291A
? CURRENT FILING DATE: 1997-06-30
? EARLIER APPLICATION NUMBER: 08/616,693
? EARLIER FILING DATE: 1997-03-13
? NUMBER OF SEQ ID NOS: 55
? SOFTWARE: Patentln Ver. 2.0
? SEQ ID NO 55
? LENGTH: 846
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-08-885-291-55

```

[illegible][illegible]

```

RESULT      4
US-09-107-847-2
; Sequence 2, Application US/09107847
; Patent No. 6100062
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID
; APPLICANT: MICHALOVICH, DAVID
; TITLE OF INVENTION: NOVEL USE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ralner & Prestlia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,847
; FILING DATE: 30-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 97304996.8
; FILING DATE: 08-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestlia, Paul F.
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-30003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 846 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-107-847-2

```

[illegible]

APPLICANT: Diederich, Robert J.
 APPLICANT: Xu, Tian
 TITLE OF INVENTION: DELTA PROTEINS, NUCLEIC ACIDS, AND
 TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESS: PENNIE & EDMONDS
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/185,432
 FILING DATE: 21-JAN-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Mastrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7326-006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2703 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-185-432-19

Query Match 3.7%; Score 276.5; DB 1; Length 2703;
 Best Local Similarity 19.1%; Pred. No. 3e-11;
 Matches 208; Conservative 147; Mismatches 383; Indels 353; Gaps 45;
 QY 455 QNNNYGIMNSPHRGSL--APRQONIMSPRKGSPKASHQFVYAGVHSMASGN 512
 DB 1801 RNLKQVAMQSGVGGAGHMSDESDMLPKQRS-----DVSQV-----GLGN 1846
 QY 513 TGNHSSSSLSALQAISEGVTSL-----LSTLSPGPKLDNSPMNITQPSKVSNO 566
 DB 1847 NGCYASHTVSEYEADQVMSAHLDDVYRAIMTPRAHQDGKRI-----DVD 1896
 QY 567 SKSPGLGFCQDNPVSSMCSNSRDHLSDESKES-----VEGAENQRPLESKG 617
 DB 1897 ARGP-----CGLTPLMAAIVAGGGLDGEDJENNEDSTAYISDLAAGAE----- 1942
 QY 618 HKLLDLTLSSDDRGSS-----LTNSPLDSSCKESVSYSPTSGVSS 662
 DB 1943 -----LNTATMKTGTSJHLARFARADAKRLFHAGADANCQDWTGRTPLAAVAAD 1995
 QY 663 TSGCVSTSMHGSILQEKRIHLKLLONGNSPAEVAK-----TTAQA--TGKD 709
 DB 1996 AMG-----VQILLNRNTNLNARHMDGTTPLILARLAIEGWEDLLTADADINAAD 2048
 QY 710 TSSITSCGDGVVQEOQLSPKKENNALRYLLDRDDPSALSKELOPYEGVDNKMSSOC 769
 DB 2049 NSGKTALHMAAVNTEAV-----NILLMHANRD----- 2078
 QY 770 TSSITPSSQEKPKITETSEBSGD-----LNDLAILGLDITSSDYNN 815
 DB 2079 -----AODKDETPFLAARESEYECALKALDNFANREITDHMDRLPRDVASERLHHD 2131
 QY 816 SISNSGSHLTQOQVFGCTSLGKSSQSVQIRPYNRAVSLDSPVSSPPVKRISA 875

DB 2132 IYRLDHEVSRPQM-----LSMTQAMISGPPGQO--- 2163
 QY 876 FPMIPKOPML-----GPNPMADSOEYSSMGSPRNVTVTQPS-----SGDWGL 922
 DB 2164 QPOLITQPTVISANGN-----NGNNSAKOSNOTAKAKAAKAKLIGSPDNL 2215
 QY 923 --PNSKGRMEPMNSNGRGGDYNTSLPRPALG--STPTLRLNSIRGARPLYLQO 978
 DB 2216 DATGSLRRKASSKKTSAASKKANLNGLDLTGVSYPVPYPPINSAYQAAAAA 2275
 QY 979 QOML-QMRPEIPMGKAN--PYGQAAASNOLGSPDGMLSMEQVSHGTQNPRLN 1034
 DB 2276 AAMSHELEGSPPVCGMGCNLPSPYDTSMTXSNMAAP--LANGNPRTGAKOPP--SY 2328
 QY 1035 DDLVPPSNLEGSDEKALLDLHTLSTNDATGLETIDALIP--ELVNOGALPEK 1091
 DB 2329 EDICKNAQMSQSLG-----NGIDMTKLDYAVASMSPPQOELLN--GQGLGN 2375
 QY 1092 QDAFO-QGEAAVMMIDQAGLYGOTYPAQGPMPMGGFHLOGSPSFMNMNMNOGNFPL 1150
 DB 2376 GNGORNGVGVLPGLGCGMGLSGAGNSRE-----QGLSPYSNQSPPHYSQSLAL 2430
 QY 1151 QGMHPRANIRPRTNP-----KOLRMOLQORLQGOOPLNQ----- 1186
 DB 2431 -SPRAYLSPSPAKSLPSLPTSPTHIQMRHATQKQFGSNLSLIGANGGVGCGG 2489
 QY 1187 -----SRQALEKMEPTAGGAVMRPMQPOGFLNMQVAKRSBELLSHFQ 1236
 DB 2490 GGGGVGQGPQNSPVSIGLISPTGSDMI--MLAPQSSKNSAIMQITSPQ-----Q 2539
 QY 1237 ORVAMMMQO 1296
 DB 2540 QO-----OQO 2388
 QY 1297 PQGFYQPNYMGQOPPARGVSSPPNAMS--SRMGPSQ--PMQHPOAASITYO--- 1349
 DB 2589 -----DSFHSGM-----NPSIOSMSGSSPSTNMLSPSSQHNOQA-FYQYLT 2631
 QY 1350 -SSEMG-----WPSGNIARNSFSQ-----Q 1370
 DB 2632 PSSQSHSGHTPOHLVOTLDSYTPSPSPGHWSSSSPSRNSDSEGVSPAANNLYISG 2691
 QY 1371 QFAHOGNAVY 1381
 DB 2692 HQANKSEALY 2702
 RESULT 6
 US-08-227-536-2
 ; Sequence 2, Application US/08227536
 ; Patent No. 5658784
 ; GENERAL INFORMATION:
 ; APPLICANT: Ecken, Richard
 ; APPLICANT: Ecken, Mark
 ; APPLICANT: Livingston, David
 ; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
 ; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Weingarten, Schurjin, Gagnedlin & Hayes
 ; STREET: Ten Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: US
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,536
 ; FILING DATE: 14-APR-1994

```

; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen A.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: DFCI-308XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-227-536-2

```

Query Match 3.5%; Score 259; DB 1; Length 2414;

Best Local Similarity 21.3%; Pred. No. 5e-10; Mismatches 335; Indels 304; Gaps 38;

```

Matches 194; Conserved 79; Mismatches 335; Indels 304; Gaps 38;

QY 575 CDQNPVSSMCOSSNRDHLSDKESKESSEVEGAENQRPLESKGHKLLQLLTCSDDRGH 634
DB 1686 CEDYDLCITCYMTKNKNDHMEKLGLDDESNNQAAATQSPGDSRRLSIORCISLVHA 1745
QY 635 SSLTNSPLD-SSCKESSVSVTSPGVSSTSGCVSSTNM-----HGSLLQE----- 680
DB 1746 COCRNANCSLPSCKQKKRVYQHTKCKRKTNGCCPICQOLIALCCYHAKHCQENKCPVPF 1805
QY 681 ----KHRIHLKLLONGNSPAEYAK---ITAOATGKDTSSITSGDGNVVKQOLSPKKKE 733
DB 1806 CLNITOKKLQOOLQHLQOQMLRRRMAASMORTG-----VVGQOQGLP----- 1848
QY 734 NNALLRYLLDRDDPSDLSKELQVEGYDNKMSQCTSTIPSSQEKDPKIKTETSEEG 793
DB 1849 -----SPRPA-----TPTPTGQGPPTTQPTQPTSGP- 1875
QY 794 SDDLNLDAIILDLTSSDPTNNSSISNGSHLGTQOVFGT--NSLGLKSSQSVQSTIRPP 851
DB 1876 -----SDDLNLDAIILDLTSSDPTNNSSISNGSHLGTQOVFGT--NSLGLKSSQSVQSTIRPP 1885
QY 852 YNRASVLDSPVSGS-----SPPVKNISAFPMILPKOP-----MLGKNPRMDSQENYSS 901
DB 1886 YLPRQOAGPVSOGRKAGAYTPTPTQTPQRPPLRPPRAVENAIOIAETQROMA-- 1943
QY 902 MCGPNRNTVTQTSSGDMGLPNSKAGMEPMNUSMGRPGDDYNTSLRPLAGSIPPL 961
DB 1944 -----HVOIFQRPQHQ--MRPMP--MAPGMN-----PPMTRG--PSG 1978
QY 962 PLRSNSIPARVLOQO-----QMLQM-----REGELPMGMGANPYGOAASNL 1007
DB 1979 HLE---PGMGPTGMQOQPPMSOGGLPQPOLQSGMPPRMMSVAGHGPPLNNAP---QP 2031
QY 1008 GSPMDGMLMEQVSHCTORPLRLNSLDDLVPSPNLQESQEBRALLDQLHTLLSTDTAT 1067
DB 2032 GL---GQVGISPLKRGTVSQALQNLRLTRSPSSPLQOQO-----VLSILH----- 2075
QY 1068 GLEEIDRALGIELVYNQGALEPKODAFQGEAAVMMDDKAGLYGQTYPAQGP----- 1120
DB 2076 -----ANFOLL-----AATIKQRAAKYANSNPQPIPGQGMPO 2108
QY 1121 -----PMGGFHLQGSPPFSNM-----MNOQNGCNFLQGMHPRANIM 1160
DB 2109 GQPGILOPRTMPQOQGVHNSPRAMQNNMPQAGYORAGLPOOQPOOQLOLPPGMSPPAOQM 2168
QY 1161 RRTGT--PKOLMLOLQORLOGOFLNOSRQALELKMENTAGAAVWRMMQPOGFLMA 1219
DB 2169 MNHNHTMPSQFR-----DLIRQOMMOQOQO-----QAGAGTIGPMANNNQFOPOQGVGTP 2220
QY 1220 QMYAQRSELLSHHFQOQRYVAMMMQOQOQO-----QOQOQOQOQOQOQOQOQOQO 1269
DB 2221 PQQRQ-----MQNHNQOMQOQGNMGQIGQLPQALGAAGASLQAYQQRLLDQOQMSPVQPN 2276

```

```

QY 1270 AFSPP-----PNVTASPSMDGLLAGPTMPQAPQPPYQ-PNYGMCQOQDPAFGRVSSPPN 1324
DB 2277 PMSPOOHMLPNOQOSEPHLOG-----QOJPNLSLQVRSRPPVPSRPSQOPRN 2324
QY 1325 AMSSSMGSPQPMQNHNPAASTY-----QSEKKGWPSGLANSSPSQOQFHHQGNP 1378
DB 2325 SSPPRMPQPPSPHNHVSFQTSPPHGLVAQAQNPMEQGHFASPDQNSMLSQL-----ASNP 2380
QY 1379 AVYSVHNNGSS 1390
DB 2381 ---GMANLHGAS 2389

```

RESULT 7

PCT-US95-04682-2

Sequence 2, Application PC/TUS9504682

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESS: Weingarten, Schurglin, Gagnebin & Hayes

STREET: Ten Post Office Square

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04682

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/227,536

FILING DATE: 14-April-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Holliday C. Heine, Ph.D.

REGISTRATION NUMBER: 34,346

REFERENCE/DOCKET NUMBER: DFCI-308Xq999

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-2290

TELEFAX: (617) 451-0313

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2414 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-04682-2

Query Match

Best Local Similarity 21.3%; Pred. No. 5e-10;

Matches 194; Conserved 79; Mismatches 335; Indels 304; Gaps 38;

```

QY 575 CDQNPVSSMCOSSNRDHLSDKESKESSEVEGAENQRPLESKGHKLLQLLTCSDDRGH 634
DB 1686 CEDYDLCITCYMTKNKNDHMEKLGLDDESNNQAAATQSPGDSRRLSIORCISLVHA 1745
QY 635 SSLTNSPLD-SSCKESSVSVTSPGVSSTSGCVSSTNM-----HGSLLQE----- 680
DB 1746 COCRNANCSLPSCKQKKRVYQHTKCKRKTNGCCPICQOLIALCCYHAKHCQENKCPVPF 1805
QY 681 ----KHRIHLKLLONGNSPAEYAK---ITAOATGKDTSSITSGDGNVVKQOLSPKKKE 733
DB 1806 CLNITOKKLQOOLQHLQOQMLRRRMAASMORTG-----VVGQOQGLP----- 1848

```

```

QY 734 NNALLRYLDRDDSPALSKELQPOVEGVNKKMSOCTSTIPSSOEKDKIKETISEEG 793
DB 1849 -----SPTPA-----TPTPTPGQQTPTTQTPTPTSP 1875
QY 794 SGDLNLDAILGLTSSDFYNNSSISNGSHLTKOOVFGT--NSLGKSSQSVQSTIRP 831
DB 1876 -----QTPPNSM-----PP 1885
QY 852 YNRAVSLDSPSVGS-----SPYKKNISAFPMI.PKOP-----MLGPRMMSOENYSS 901
DB 1886 YLPRTOAGAVSOGKAGAVTPTPTPTAOPPLPGPPYAVEMAOIOTALETOROMA-- 1943
QY 902 MGSPPNNVTYQTPTSSGDMGLPNSKAGRMEMNSMGRPGCDYNTSLPPALGGSITPL 961
DB 1944 -----HVOIFQRIHQ--MPPMTF--MAPMGMN-----PPMTRG--PSG 1978
QY 962 PLRSNSTPGARVYLOOQO-----OMLOM--RPEIIPMGANPYGOAASNOL 1007
DB 1979 HLE-----PGMGPTMOOQOPWMSOGGLPPOQLOLSCMPRPMMSVAQHGPPLNAP--QP 2031
QY 1008 GSWPDGMSMEVSHGTONPRLRLNSLDDLVGPPSNLEGSDEKALLDOLHTLSNTDAR 1067
DB 2032 GL-----GQVGSIPKPGTVSQOALQNLRLTLPSSPLQOQO-----VLSTLH----- 2075
QY 1068 GLBEIDRALGIPETLVNOGALPEPKDAFQGEAAYVMDOKAGLYGOTYPAQGP----- 1120
DB 2076 -----ANPOL-----AATIKOAAKTAANSNPPTFGQGPMPQ 2108
QY 1121 -----PMOGFHLQGSPPSFSM-----NMOMNOGNFPLQGMHPRANIM 1160
DB 2109 GQGLDPRTPRGQGGVNSNPAMNMNMQAGVQAGLPPQOQOQDPRMGSMSPQOQM 2168
QY 1161 RPTNT--PKOLMLOOQOGLQOQFLNDSKALBKEMNPAGGAAYVRPMQOQGLNA 1219
DB 2169 NMNMTWPSQFR-----DIFRQOQMOOQOQ-----OGAGPGICPGMANHNOFOQPOQGYR 2220
QY 1220 QWVAORSRELLSHHFRQORVAMMMQOQOQO-----QOQOQOQOQOQOQOQOQO 1269
DB 2221 PQRQR-----MOHMOQMOQNGMOIGOLPOLGAEAGSLAQOQLLOOQMGSPYOPN 2276
QY 1270 AFSPR-----PNVTASPSMDGLLAGPTPQAPPOQFPYQ--PNYMGQOPDPAFQVSSPPN 1324
DB 2277 PMSPOHMLPNOASPLHOG-----QOIPNSLSNOYRSPQVPVSPRQSGPRH 2324
QY 1325 AMMSRRGQSPQNMOMHPOAASTY-----QSEBMKGWPSGMLARNSFQOQOFAHGNP 1378
DB 2325 SSFSPRQOPQSPHNSPOTSSPHPLVLAQAANPMEGHPASPDQNSMLSQL-----ASNP 2380
QY 1379 AVYSVMHMGSS 1390
DB 2381 ---GMANLHGAS 2389

```

```

RESULT 8
US-08-785-241-6
: Sequence 6, Application US/08785241
: Patent No. 5695963
: GENERAL INFORMATION:
: APPLICANT: McKnight, Steven L.
: APPLICANT: Russell, David W.
: APPLICANT: Tian, Hui
: TITLE OF INVENTION: Endothelial PAS Domain Protein
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 BUSH STREET, SUITE 3200
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

```

```

: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/785,241
: FILING DATE: 17-JAN-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: OSMAN, RICHARD A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: UTSD:1229
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 343-4341
: TELEFAX: (415) 343-4342
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 826 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-08-785-241-6

```

```

Query Match 3.5%; Score 258; DB 1; Length 826;
Best local similarity 19.8%; Pred. No. 1e-10;
Matches 181; Conservative 136; Mismatches 320; Indels 278; Gaps 41;

```

```

QY 1 MSGIGENLD--PLASDRKRRLPCDTPGGLTSGEKKRRREQSKYIEELAEIISANLSD 58
DB 1 MEGAGANDKKKISSEKREK-----SRDAARSRSKSEVEYELAH-----QLPL 46
QY 59 IDNFNKPDKCALIKEVYR--QIRQIKEGKTTISNDDVQKADVSSFGVVIDKDSGL 116
DB 47 PHNVSSHLDKASVAVRLTISYLRVRKLLDAG-DIDIEDM-----KAQMNCPR 91
QY 117 LLOALDELFFVNRKANIVVSENVTOYLOKQEDLVNTSVNILHEEDKDKFLNLPKS 176
DB 92 YLKALDGFVAVLDDGDMITISDNVKNKMGTLQFELTGHVFPFTPHCHIEKREMLTHR 131
QY 177 TVNG-VSWTNEPOROKS-----HTENC-----RMILKTPHDLIEDINASPEN 217
DB 152 --NGLVKKGEQNTQRFELFMKCTLTSGRTMNIKSAWKVLHCTGHIHYVDTSNQP- 208
QY 218 RQRYETMOCALSOPRAMEEGEDLOSCMTCVARITTGRTSPNE-----SPTTRH 271
DB 209 -----QCGYKKRP-----MTCLVLCEPI-----PPHSNIEPLDSKTELSRH 246
QY 272 DLGKVVNIIDTNSLRSSMRPGF--EDIIIRCIORFESLNDGQSMQSRHAYQEAYLNGHAE 329
DB 247 SLDMKFSYCD--RITELMGYEDELGSIYIEYHALDSHLTKTH--DMFTKGQVT 301
QY 330 TPVYRFSLADGTVTAQTSKSLFRNPVTNDRHGFSVTHFLQRBQNGYRPNPNVGGQIRP 389
DB 302 TGOYRMLAKRGYVWVETQATVIYNTKNSQPOCIVCVNY----- 341
QY 390 PMACGNSVCGM-----SMSPNOGLQMPSSRAYGLADPSTTGQMSGARVYG 435
DB 342 -----VSGIIOHDLIFSLQOETCEVLKVESSDMKMTQLEFKVESEDTSSFLDKLKE 333
QY 436 SSNIASLTTPGQMSPSSYONNNYGLNMS-----PPHSSPGI-APNOO-----NIM 481
DB 394 PDALITLAPAG--DTIISLDFGSDTETDQOQLEEVLYLNDVMLPSNREKLQNTLNA 449
QY 482 ISPRNRSFKIASHQSPVAGVHSPMASSGNTGNHSESSSSLSALOAISEGCVTSLSTL 541
DB 450 MSP-----LPTAETPKPLRSSAD-----PALNOEVALKL----- 478
QY 542 SSPGPKLNSPNNINQ-----SKVSNDSKSP--LIGFYQONVSSMCQ 586
DB 479 -EPNPE--SLEISFTWPOIQODTPSPSDGSTQSSPEPSPSEYCYVDSDWNEFKLE 534
QY 587 SNRDLHLSKESSESSVEGAENQGPLESKGRKHLQQLT--CSSDD-----RGHSSLTNS 640

```

```

Db      535  LVEKLFADTEAK-----NPESTODTDLEMLAPYIPMDDFQLARSFDL--S 583
Oy      641  PLDS-----CHSSVYSPSPGCVSSSPGCVSSNNHGSLQKXKHLH-KLLQNG 692
Db      562  PLESSASPEBASPOSTVYFQQTQIQDEPTIATTTATTDKVTYKDMEDIKILIAS 641
Oy      693  NSPAEVAKITAQATG---KDTSSITS---CGDG-----NVV 722
Db      642  PSPFHKHETTSATSSPYRDIQSTASPNRAGKVIIEQTEKSHRSPNSVALSORITV 701
Oy      723  KOEQLSPK-----KKNALL-----RYLLRDDPSDALSKELOPOVEGV--- 762
Db      702  PEEELNPTIILAQNAQRKRMEHGSLFOAVGIOTLLQCPDDHNAATTSLSMKRVKCKSS 761
Oy      763  -DNKMSOCTSTIPS 776
Db      762  EONGMEQKTTILIPS 776

```

```

1      RESULT          9
2      US-08-480-473B-2
3      ; Sequence 2, Application US/08480473B
4      ; Patent No. 5882914
5      ;
6      ; GENERAL INFORMATION:
7      ; APPLICANT: Semenza, Gregg L.
8      ; TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
9      ; NUMBER OF SEQUENCES: 64
10     ; CORRESPONDENCE ADDRESS:
11     ; ADDRESSEE: Fish & Richardson P.C.
12     ; STREET: 4225 Executive Square, Suite 1400
13     ; CITY: La Jolla
14     ; STATE: CA
15     ; COUNTRY: USA
16     ; ZIP: 92037
17     ;
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Floppy disk
20     ; COMPUTER: IBM PC compatible
21     ; OPERATING SYSTEM: PC-DOS/MS-DOS
22     ; SOFTWARE: Patentin Release #1.0, Version #1.30
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/08/480,473B
25     ; FILING DATE: 06-JUN-1995
26     ; CLASSIFICATION: 514
27     ;
28     ; ATTORNEY/AGENT INFORMATION:
29     ; NAME: Halle, Lisa A.
30     ; REGISTRATION NUMBER: 38,347
31     ; REFERENCE/DOCKET NUMBER: 07265/053001
32     ; TELECOMMUNICATION INFORMATION:
33     ; TELEPHONE: 619/678-5070
34     ; TELEFAX: 619/678-5099
35     ;
36     ; INFORMATION FOR SEQ ID NO: 2:
37     ; SEQUENCE CHARACTERISTICS:
38     ; LENGTH: 826 amino acids
39     ; TYPE: amino acid
40     ; TOPOLOGY: linear
41     ;
42     ; MOLECULE TYPE: protein
43     ;
44     ; US-08-480-473B-2

```

Query Match	3.5%	Score 258;	DB 2;	Length 826;
Best Local Similarity	19.8%	Pred. No. 1e-10;		
Matches 181;	Conservative 166;	Mismatches 320;	Indels 278;	Gaps 41

```
QY      1 MSGGENID--PLASDSKRRLPDDTGGGJLTCGSGERREOREKYIELLAELISANLSD 58
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1 MEGAGGANDKKRKSERRKEK-----SRDAARSKRSSEVPEIAH---QLPL 46
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      59 IDNFNVKPKDCAILKETVR-QIROIKECKTTISNDDVOAKADVSSTGGGVIDKDSLGPL 116
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      47 PHNVSSHLDKASVRLTISYLVRKLDDAG-DLDEIDDM-----KAQNCF 91
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      117 LLAGALDFLEFVNREAAIVAFSVANTQYLQKQDDLVTNTSYNLLHEEDRKDFLNLPKS 176
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

Dd	92	YLKALDGEVWVLTDGDGDIYISDNVNNKYMGLTQELGHSVFDPTHHCDEHREKREMLTHR	151
Qy	177	TVNG-VSMTNRPQOKS-----HTFNC-----RMLKTPHDLLEDINASPREM	217
Dd	152	--NGLVKRGKQDNQORSFELRMKCTSLTRSGTNNIKSAYMKVLHCTGHINUYDITNSQP-	208
Qy	218	RORYETMOCFALISOPRAMEEGEDLOSOMICVARIRITGTERFPSPNDE-----SFYTRH	271
Dd	209	-----QCGKTRP-----MCLVLCRPT-----PHPSNIEIFLDKTFISRH	246
Qy	272	DLSGKVNIDITNSLRSSMRPGF--EDIIIRCIORFFSLNDGQSNQSRHYQOAYLNGHAE	329
Dd	247	SIDMKFSYCYDE--RIFELMGYEEBELGRSEIYEYHALDSHDLTKTHN--DMETKGQVY	301
Qy	330	TPVYVFSIADGTIYTAQCKSKLFERNPTYNDBHGVSHTHFLQREDONGIRPNPNVGGCIRP	369
Dd	302	TGQYMLAKRGGYVWVEQTAVIYNTKNSQPCITCVNY-----	341
Qy	390	PMAGCNSSVGCN-----SMSPROGLOMPSSRAYGLADSTGOMSGARYG	435
Dd	342	-----VSGIQLHDILFSIQOTECVYAKRPVRESSDMKMTQLFTEVEEDTSSLEFDLAKR	393
Qy	436	SSNIASLTPEGKMOSSPSYQNNNTGLNKS-----BPHSGPL-APNOQ---NIM	481
Dd	394	PDALITLLPAPAG---DPIISLDGFSNDTETDDQOLEEVLRYNDVMLPSPNEKLOINILA	449
Qy	482	ISPRRGRSPKIASHQFSPVAGVHSPMASGGTGNHSSSSSLASLOALISGCVGSLISTL	541
Dd	450	MSP-----LPTAETPRPLRSSAD-----PALNOEVALKLT--	478
Qy	542	SSPGKILDNSENMTNTP-----SKVSNODSKSP--LGFYCOQNPVSSSMQ	586
Dd	479	-EPNE--SLESTPTMQIQDQTPSPSDGSTRQSPSPNPSBYCYVYSDWVNEKLE	534
Qy	587	SNSRPHLSDKESKESVGAENQKPLESKHKKLLDOLL--CSSDP---RGHSSTJNS	640
Dd	535	LYEKLEAFDETRAK-----NPESTQOTDULEMLAPYIPMDDEDFOLRSFDOL--S	581
Qy	641	PLDSS-----CKRESSYVTPSPGVSYSTSGVSSNNHGSLSQKXKHLH--KILONG	692
Dd	562	PLESSASBPESASPOSTIVVPOQTOIOEPTANATTTATTDDELKVTYTKDRMEDIKILIAS	641
Qy	693	NSPAEVANTQAQATG--KDRYSITS-----CGDG-----NVV	722
Dd	642	PSPTIHKHETTSATSSPYRDOQSTRASPNRAGKCVIQTOKESHPRSPNIVSVALSQRFTV	701
Qy	723	KOQLSPK-----KENNALL-----RYLDRDDPSALSKELQPOVEGY--	762
Dd	702	PEEELNPKYIILONNOKRKKMEHSGSLEQAVGICTITLLOQPDHNAATTSLSMKRYKCKSS	761
Qy	763	-DNKSQCTSSITPS	776
Dd	762	EONGMEQKTIILIPS	776

```

10. RESULT 10
11. US-08-915-213-2
12. : Sequence 2, Application US/08915213
13. : Patent No. 6020462
14. :
15. : GENERAL INFORMATION:
16. :
17. : APPLICANT: Semenza, Gregg L.
18. :
19. : TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
20. :
21. : NUMBER OF SEQUENCES: 64
22. :
23. : CORRESPONDENCE ADDRESS:
24. :
25. : ADDRESSEE: Fish & Richardson P.C.
26. :
27. : STREET: 4225 Executive Square, Suite 1400
28. :
29. : CITY: La Jolla
30. :
31. : STATE: CA
32. :
33. : COUNTRY: USA
34. :
35. : ZIP: 92037
36. :
37. : COMPUTER READABLE FORM:
38. :
39. : MEDIUM TYPE: Floppy disk
40. :
41. : COMPUTER: IBM PC compatible
42. :
43. :

```


OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915, 213
FILING DATE: 20-AUG-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/480,473
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/053001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 826 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-915-213-2

Query Match 3.5%; Score 258; DB 3; Length 826;
Best Local Similarity 19.8%; Pred. No. 1e-10; Indels 278; Gaps 41;
Matches 181; Conservative 136; Mismatches 320;

QY 1 MSGIGENLD--PLASDRKRRLPCDTPGGLTSCGKRRREDSKYIEELAEIISANLSD 58
DB 1 MEGAGANDKKKISSERRK-----SDAARSRSKSEVEFYELAH---QLPL 46
QY 59 IDNFNVKPKCAILKEFVR--QIRQKEGKTTISNDDVQKADVSSTGGVIDKDSIGPL 116
DB 47 PHNVSSHLDKASVWRLLTISLYRKRLLDAG--DIDIEDM-----KAQMNCF 91
QY 117 LIALDGFLEVVNREANIVFSENVTOYLOKQEDLVNTSVNLIHEEDRKDFLKNLPKS 176
DB 92 YLKALDGFVAVLTDGDMITISDNVKNKMGTLQFELGHSVDFPTPCHEERKREMLTIR 151
QY 177 TVNG-VSWTNEPOROKS-----HTFNC-----RMLAKTPHDLIEDINASPEM 217
DB 152 --NGLVKKGEQNTORSFLRMKCTLSRGRTNINISATKVLHCGHIVYDTNSQP- 208
QY 218 RQRYETMOCFALSOPRAMEEGEDLOSMTCVARRITTGERTPSNPE-----SITTRH 271
DB 209 -----QCGYKKRP-----MTCVLVLEPI-----PHPSNIEIPLDSTFLSRH 246
QY 272 DLSGKVVNIDTNSLRSSMRPGF--EDIRRCIORFSLNDGQSMQKRHYOEAYLNGHAE 329
DB 247 SLDMKFSYDCE--RITELMGTEPEELIGRSITYEYHALDSHLTKTH--DMFTGQYV 301
QY 330 TPVYRSLADGTLVTAQTSKLEFRNPVTNDRHGFVSTHFLQROQNGYRPNPNVGGGIRP 389
DB 302 TGGYRMLAKRGYVWETQATVITYNTRKNSQPOCIVCNV----- 341
QY 390 PMAGCNSSVCGM-----SMSPNGQLQMPSSRAVYGLADPSTTGQMSGARVGG 435
DB 342 -----VSGIIOHDLIFSLQOTECVCLRVESSDMKMTQLFTKVSSEDSISFLDKLKE 393
QY 436 SSNIASLTGPGCMQSPSSYQNNNYGLNMS-----PRHGSPIG--APNOQ-----NIM 481
DB 394 PDALITLLAPAG---DTIISLDGFSNDTETDQQLAEVPLVDVMLPSNEELONINILA 449
QY 482 ISFRNKGSPKIASHOPSPVAGVHSPMASSGNTGNHSSSSSLALQAISEGVTSLSTL 541
DB 450 MSP-----LPAETPKPLRSSAD-----PALNGEVALKL----- 478
QY 542 SSFGPKLDSNNNTQIP-----SKVSNODSKSP--LGFYCDQNPVESSMCQ 586
DB 479 -EPNPE---SLEISFTMPOIQDOTPSPSDGSTRQSSPEPNSPSEYCFYVDSDMVNEFKLE 534

QY 587 SNRDHLDSKESKESSEVEGAENQRPLESKHKKLLQLLT--CSSPD-----RGHSSLTNS 640
DB 535 LVKLELAEDTEAK-----NPESTODTDLDEMLAPYIPMDDDQLKRSFDOL--S 561
QY 641 PLDS-----CKESSVATSPGVSSTSGVSSSTNMHSGSLQEKHRLH--KILONG 692
DB 582 PLESSASPEASAPQSTVTVFQQTQIQEPTANNTTATTDKLTATYKQMEDIKILLAS 641
QY 693 NSPAEVAKITAATG--KDTSSITS---CGDG-----NVV 722
DB 642 PSPTHIKETTSATSSPYRDTQSRTASPNRAGVGIEQTEKSHRSPNVLSVALSORTVY 701
QY 723 KQSLSPK-----KENNALL-----RYLDRDPDADLSKELQPOVEGV--- 762
DB 702 PEEELNPKILLALONAOKRKMEHDSLPQAVGIGTLQDPDDHAATYSLSWKRVKCKSS 761
QY 763 -DNKMSQTSSTIPS 776
DB 762 EQNGMEQKTIILIPS 776

RESULT 11
US-09-148-547-2
Sequence 2, Application US/09148547
Patent No. 6124131

GENERAL INFORMATION:
APPLICANT: Semenza, Gregg L.
TITLE OF INVENTION: Hypoxia Inducible Factor-1 and Methods of Use
FILE REFERENCE: 07265/151001
CURRENT APPLICATION NUMBER: US/09/148,547
CURRENT FILING DATE: 1998-08-25
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 826
TYPE: PRT
ORGANISM: Homo sapiens
US-09-148-547-2

Query Match 3.5%; Score 258; DB 4; Length 826;
Best Local Similarity 19.8%; Pred. No. 1e-10;
Matches 181; Conservative 136; Mismatches 320; Indels 278; Gaps 41;

QY 1 MSGIGENLD--PLASDRKRRLPCDTPGGLTSCGKRRREDSKYIEELAEIISANLSD 58
DB 1 MEGAGANDKKKISSERRK-----SDAARSRSKSEVEFYELAH---QLPL 46
QY 59 IDNFNVKPKCAILKEFVR--QIRQKEGKTTISNDDVQKADVSSTGGVIDKDSIGPL 116
DB 47 PHNVSSHLDKASVWRLLTISLYRKRLLDAG--DIDIEDM-----KAQMNCF 91
QY 117 LIALDGFLEVVNREANIVFSENVTOYLOKQEDLVNTSVNLIHEEDRKDFLKNLPKS 176
DB 92 YLKALDGFVAVLTDGDMITISDNVKNKMGTLQFELGHSVDFPTPCHEERKREMLTIR 151
QY 177 TVNG-VSWTNEPOROKS-----HTFNC-----RMLAKTPHDLIEDINASPEM 217
DB 152 --NGLVKKGEQNTORSFLRMKCTLSRGRTNINISATKVLHCGHIVYDTNSQP- 208
QY 218 RQRYETMOCFALSOPRAMEEGEDLOSMTCVARRITTGERTPSNPE-----SITTRH 271
DB 209 -----QCGYKKRP-----MTCVLVLEPI-----PHPSNIEIPLDSTFLSRH 246
QY 272 DLSGKVVNIDTNSLRSSMRPGF--EDIRRCIORFSLNDGQSMQKRHYOEAYLNGHAE 329
DB 247 SLDMKFSYDCE--RITELMGTEPEELIGRSITYEYHALDSHLTKTH--DMFTGQYV 301
QY 330 TPVYRSLADGTLVTAQTSKLEFRNPVTNDRHGFVSTHFLQROQNGYRPNPNVGGGIRP 389
DB 302 TGGYRMLAKRGYVWETQATVITYNTRKNSQPOCIVCNV----- 341
QY 390 PMAGCNSSVCGM-----SMSPNGQLQMPSSRAVYGLADPSTTGQMSGARVGG 435

```

Db 342 -----VSGIIQHDLIFSLOQTECVLKPVSSDMKMTQLETKVESEDTSSLFDKKE 393
Qy 436 SSNIASLTTPGPMQSPSSYONNNYGLNMS-----PRHSGPL-APNOQ-----NIM 481
Db 394 PDALTLAPAAQ-----DTIISLDFGSNDTETDDQLEEVPLVNDVWLPSPNKKLQINILA 449
Qy 482 ISPRNRGSKTASHQSPVAGVHSPMASSGNTGNHSFSSSLALQAISEGVGTSLSLTL 541
Db 450 MSP-----LPTAETPKPLRSSAD-----PALNOEVALKL----- 478
Qy 542 SSPGPKLDNSPMNITOP-----SKVSNODSKSP--LGFYCDONPVSSMCQ 586
Db 479 -EPNPE-----SLELFTMQIDQTPSPSDGSTROSSPEPNSSEYCFYVDSMVNEFKLE 534
Qy 587 SNSRDHLSDKESKESSEVEGAENORGPLESKGHKLLQILT--CSSDD-----RGHSLSLNS 640
Db 535 LVEKLEAEDTEAK-----NPFSTQDTDLDEMLAPYIPDDDDPQLKSPQL--S 581
Qy 641 PLDSS-----CKSSSVTSPSGVSSSTSGCVSSTSMHGSLLQEKHRIH-KLONG 692
Db 562 PLESSASPEASPOSTVTVFOQTQIOEPTANATTTATDELKTVTYKDRMEDIKILIAS 641
Qy 693 NSPAEVAKITTAQATG---KDTSSITS---CGDG-----NVY 722
Db 642 PSPTHIHETTSATSSPYRDQSRITASPNRACKGVIETEKSHRSPNVLSVALSORTTV 701
Qy 723 KOEQLSPK-----KKENNAL-----RYLLDRDDPSDALSKELOPOEGV--- 762
Db 702 PEEELNPKILALQNAQRKKMEHDSLFQAVGIGTLQOPDHAATTSLSMKRVKGCSS 761
Qy 763 -DNKMSQCTSSITPS 776
Db 762 EQNGMEQKTIILIPS 776

RESULT 12
PCT-US96-10251-2
: Sequence 2, Application PC/TUS9610251
: GENERAL INFORMATION:
: APPLICANT: The Johns Hopkins University School of Medicine
: TITLE OF INVENTION: HYPOXIA INDUCIBLE FACTOR-1 AND METHOD OF USE
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESS: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/10251
: FILING DATE: 06-JUN-1996
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Haile, Lisa A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 07265/033W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5099
: TELEFAX: 619/678-5070
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 826 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US96-10251-2
```

```

Query Match 3.5%; Score 258; DB 5; Length 826;
Best Local Similarity 19.8%; Pred. No. 1e-10;
Matches 181; Conservative 136; Mismatches 320; Indels 278; Gaps 41;

Qy 1 MSGLEND--PLASDRKRRLPCDTPGGGLTCSSEKRRREDESKYIELLAILSANLSD 58
Db 1 MEGAGAGNDKKKISSERRREK-----SRDAARSRKSESEVFYELAH---QLPL 46
Qy 59 IDNFNVKPKCAILKEAYR--OIROIKEGKTISSDDVQKADVSTSGGVIDKDSGLPL 116
Db 47 PHNVSSHLDKASVMKLTLSYLRVRLKLDAG-DLIDIEDD-----KQNMCF 91
Qy 117 LQALDGLFLEVNRANIVSEANTQYLYQKQEDLVNTSVYNIILHEEDRKDFLKNLPS 176
Db 92 YLKALDGFVWVLTDDGDMITYSIDNVNKKYMGTLQFELTGHVSVDPLPHPDHEMRMLNHR 151
Qy 177 TYNG-VSMTNEPQOKS-----HTFNC-----RMKMTPRDILEDINASPDM 217
Db 152 --NGLVKKGKEQNTQRSEFLRMRKCTLTSRGRTMNKSAWKVLCGHIIHYDTSNQP- 208
Qy 218 RQRYETMOCFALSOPTAMEBEGEDLQSCMICVARRITTGERTFSPNPE-----SFTRH 271
Db 209 -----QCGYKRP-----MTCLVLICEPT-----PHPSNIEPLDLSKTFELSRH 246
Qy 272 DLSGKVNIDTNSLRSMRPF--EDIIIRCIQRPESLNDGQSWSQKRHYQEAFLNGHAE 329
Db 247 SLDMKFVYCD--RITELMGYPEPELLGRSLEYEYHALDSDHLKTHH--DMFTKGGVT 301
Qy 330 TPVYRPSLADGIYTAQTKSKLEFRNPVINDRNGFVSTHFLQEQNGYRPNRPVQGGIRP 389
Db 302 TGOYRMLAKRGQYVWVETQATVYNTKNSQPOCIYCVVY----- 341
Qy 390 PMAGCNSSVYGM-----SMSPNGQLMPSSRAYGLADPSTTGMSGARYGG 435
Db 342 -----VSGIIQHDLIFSLOQTECVLKPVSSDMKMTQLETKVESEDTSSLFDKKE 393
Qy 436 SSNIASLTTPGPMQSPSSYONNNYGLNMS-----PRHSGPL-APNOQ-----NIM 481
Db 394 PDALTLAPAAQ-----DTIISLDFGSNDTETDDQLEEVPLVNDVWLPSPNKKLQINILA 449
Qy 482 ISPRNRGSKTASHQSPVAGVHSPMASSGNTGNHSFSSSLALQAISEGVGTSLSLTL 541
Db 450 MSP-----LPTAETPKPLRSSAD-----PALNOEVALKL----- 478
Qy 542 SSPGPKLDNSPMNITOP-----SKVSNODSKSP--LGFYCDONPVSSMCQ 586
Db 479 -EPNPE-----SLELFTMQIDQTPSPSDGSTROSSPEPNSSEYCFYVDSMVNEFKLE 534
Qy 587 SNSRDHLSDKESKESSEVEGAENORGPLESKGHKLLQILT--CSSDD-----RGHSLSLNS 640
Db 535 LVEKLEAEDTEAK-----NPFSTQDTDLDEMLAPYIPDDDDPQLKSPQL--S 581
Qy 641 PLDSS-----CKSSSVTSPSGVSSSTSGCVSSTSMHGSLLQEKHRIH-KLONG 692
Db 562 PLESSASPEASPOSTVTVFOQTQIOEPTANATTTATDELKTVTYKDRMEDIKILIAS 641
Qy 693 NSPAEVAKITTAQATG---KDTSSITS---CGDG-----NVY 722
Db 642 PSPTHIHETTSATSSPYRDQSRITASPNRACKGVIETEKSHRSPNVLSVALSORTTV 701
Qy 723 KOEQLSPK-----KKENNAL-----RYLLDRDDPSDALSKELOPOEGV--- 762
Db 702 PEEELNPKILALQNAQRKKMEHDSLFQAVGIGTLQOPDHAATTSLSMKRVKGCSS 761
Qy 763 -DNKMSQCTSSITPS 776
Db 762 EQNGMEQKTIILIPS 776

RESULT 13
US-08-785-310A-8
```

```

: Sequence 8, Application us/08785310A
: Patent No. 5840532
: GENERAL INFORMATION:
: APPLICANT: McKnight, Steven L.
: TITLE OF INVENTION: Neuronal PAS Domain Protein
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 BUSH STREET, SUITE 3200
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/785,310A
: FILING DATE: 21-JAN-1997
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: OSMAN, RICHARD A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: UTSD:1226
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 343-4342
: TELEFAX: (415) 343-4342
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 816 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-785-310A-8

Query Match 3.4%; Score 254.5; DB 2; Length 816;
Best Local Similarity 21.1%; Pred. No. 1.8e-10;
Matches 194; Conservative 137; Mismatches 346; Indels 241; Gaps 47;

QY 34 EKRRRQESKYEIELALISANLSDINFNKPKCALIKETVQIQIKQEGTISND 93
DB 18 EKRRRQGFNVLKELSMLEPGNTR-----KMDKTVLKEVIGLQKHNE---VSAQT 66
QY 94 DVOKADVSTGQ-GVIDKDSLGPLLQALDGLFVVRNREANIVVSENVQYLOKQEDL 152
DB 67 EI--CDIQOMKPSFLSNEFEQTQMLALDGLFYVTTDSIIVSDSTIPPLGHLRADY 124
QY 153 VNTSVYNIHEEDRKDKLPLKSTVNGVSWTNEPQOKSHT---FNCRLMKTPHDILE 209
DB 125 MDQNLNLFPEQHESEYKILSHML--VUDSPSEFLKSDNDEFLCHLLRG----- 175
QY 210 DINASPMRQRYETMOCF-----ALSOPRAMEEGEDLOSCMTCVA 250
DB 176 --SLNPKPEFTYEIKIVGNFRSNVNPSPSCNGFDNLTSLRP-CHVPLGADV--CFIATV 230
QY 251 RRII---TGERTFPSNP-ESFITRHDLSGKVVNIDTNS-----LRSSMRPGE- 294
DB 231 RLATPOLKEMCAVDELLEFTSHSLKWKFLFLDHRAPPLIGLPEVLTGS---GINY 287
QY 295 -----DIIRRCIORFSLNDGOS-----WSQKRHYQEAY--LNGHAETPVY 333
DB 288 YHIDDELLARQHQLMQLQFGKSKCCYRPLTKGQWIMLQ--THYITTYHQMNSKEPIVC 346
QY 334 RPSLADGTITYAQTAKL-FRNPVTNDRHGFSVTHPLQREONGRPNPNVYGQ---GIRP 389
DB 347 THSVSYADYRVERKQELALDEDPTEAMH-----PSAVKEDSSLEP 388
QY 390 PMAGCNSVSGMSKSPNGQLQMPSSRAYGLADPSTTGMSGARYGSSNLSLTPGGMQ 449

```

```

DB 389 P-----QPPNALDM-----GASGLPS-SPSSAS 411
QY 450 SPSSYQNNNYGLNKSPPHSPGLAQNQIMISPRNRGSPKIASHOFSVAGVHSPMAS 509
DB 412 SRSSHKSSHTAM---SEPTSTPKTMAENSTALPRPATLP-----QELPVQCL--SQAA 461
QY 510 SGNTGNHSPSSSLSS---ALQATISEGVSTLSLTLSPGKLDNSPMMNITOPSKYSND 566
DB 462 TMTPLHSSASCGLTKQLLQSLPO--TGL--QSPPAVYTOPSAQFSMEOTIK----- 510
QY 567 SKSPGLFYCDQNPVSSMCSNSRDLSDKESKESVEGAENORGPLESGHKKLLQ--- 623
DB 511 -----DQLEGRTRILQANIR-----WQOEELHKIQEOLCLVQDSNVMQMLQQA 554
QY 624 -LITGSSDDR--GHSLTNSPLDSSCKESSVYTPSGVSSSTGSGVSTSNMHSGLQ 660
DB 555 VSLSFSTQRPAAQOQLOQRPAPS--QPOLVYVTP-----LOGQLTSTQVFNQHLLRE 606
QY 681 KHRILHLKLLONGNSPAVAKITATQATKDNSSITSCGDGNVYQEOULSPKKENALLRY 740
DB 607 SNTI---SAQGRPKMSSQL-LPASGRSLSLPS-----QPSYASVLPPLGLSL 651
QY 741 LLDREDDPSALSKELQOVE-GVDNK-----MSQCTSTIPSSOEKDPKIKTETSEBGS 795
DB 652 TTIAPTPD--DSQCQSPDPGHDRLRLLSQIOPMMSGCDARP---SEVSRTRGQ 706
QY 796 DLNDLAILDGLTSSDFYNSISSNSGSLGTQOY---FQGTNSLGLKSSQVOSTIRP 851
DB 707 VKYAQSQVMFPSPDHPHTSSASAPVLLMG--QAVLHPSPAPSPPLQAOAQQO--PPP 763
QY 852 YNRVSLDSPVYSSSPP 869
DB 764 Y-----LQAPTSLSHEQP 776

RESULT 14
US-08-816-693A-53
: Sequence 53, Application US/08816693A
: Patent No. 5874241
: GENERAL INFORMATION:
: APPLICANT: Takahashi, Joseph S
: APPLICANT: Turek, Fred W
: APPLICANT: Pinto, Lawrence H
: TITLE OF INVENTION: Clock Gene and Gene Product
: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dressler, Rocky, Milnamow & Katz
: STREET: Two Prudential Plaza, Suite 4700
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/816,693A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5874241thrup, Thomas E
: REGISTRATION NUMBER: 33,268
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-616-5400
: TELEFAX: 312-616-5460
: INFORMATION FOR SEQ ID NO: 53:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 816 amino acids
: TYPE: amino acid
: STRANDEDNESS: single

```

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-816-693A-53

Query Match 3.4%; Score 254.5; DB 2; Length 816;

Best Local Similarity 20.8%; Pred. No. 1.8e-10; Matches 190; Conservative 140; Mismatches 352; Indels 231; Gaps 45;

34 EKRRROESKYIEELAEELISANLSDIDNENVPKDKALKEVROIROKEGKTIISND 93
18 EKRRROQEVVLKELSSMLPGNIR-----KMDKITVLEKVIIGLQKINE---VSAQT 66
94 DVQKADVSSTGO-GVIDKDSLGPILLQALDGLFVYNREANIVFVSENVTOYLQKQEDL 152
67 EI--CDIQQDMKPSFLSNEEFTQMLLEALDGEVIYVTTGDSITIVSDSTPLGLHPADV 124
153 VMTSVYNIHEEDRKDFLNLKSTVNGVSWTNEPQROKSHF--FNCMLMKTPTDILE 209
125 MDONLNLFLPEQHSFVYKILSSHML--VTDSPPEFLKSDNDEFLYCHLLNG----- 175
210 DINASPEMRQREYTMQCF-----ALSOPRAMEEGEDLOSCMICA 250
176 --SLNPKPEPTYEYIKFVGNFRSYNNVPSPCNGDNTLSRP-CHVPLGKV--CFIATV 230
251 RRIIT--TGERTFPSPNP-ESFTRHDLGKVVNIDTNS-----LRSSMRGFE- 294
231 RLATPOFLKEMCVADEPLEEFTSRHSLEWKFLDHRAPPIIGYLPFEVLGTS--GYNY 287
295 -----DIRRCIORFFSLNDGOS-----WSQKHGYQAY--LNGHAETPY 333
288 YIIDDELLARCHOHLMOFGKSCCYRFLTKGQWIMLQ-THYITTYHOMNSKPEFTVC 346
334 RFLSLADGITVIAQTQSKL-FRNPVTNDRHGFVSTHFLQREONGYRPNPNVGO--GIRP 389
347 THSVSYADAVRERROELALEDDPPEAMH-----PSAVKEXKSSLEP 388
390 PMAGCNSSVGVGSMSPNOGLQMPSSRAYGLADPSTTGOMSGARYGSSNINASTLPQGMQ 449
389 P-----QPFNALDM-----GASGLPS-SPSPSAS 411
450 SPSSYONNNYGLNMSPPHSGPLAPNOONIMISPRNGSPKIASHQSPVAVGSPMAS 509
412 SRSSHKSSHTAM--SEPTSTPTKLAENSTALPRATLP-----QELPVQGL--SOAA 461
510 SCNTGNHSESSSSLS--ALQAISEGVGTSLLSTLSSPGKLDNSPNNITOPSKVSNOD 566
462 TMTPTALHSSASCDLTQKLLQSLPQ--TGL--QSPAPVYQFSAQFSMOTIK----- 510
567 SKSPGLGYCDQNPVSSMQOSNRDHLSDKESKESVGEAENQGRPLESKGHKKLLQLT 626
511 -----DQLEQRTRIQANIR-----WQBELKHIQEOCLCYODSNVQMFLOQPA 554
627 CSSDDRGHSLNPLSDSCKESSSVTSPS-GVSSSTSGVSTSNMHGSLDEKXRL 685
555 VS--LSFESSIORPAAQOOLQRPAPASPOLVNTMPLQGTITQVTHQHLRESNVI- 610
686 KHLLONGNSPAVAKITTAQATGKTJSSITSCGDGVNYYKQEQLSPKKKENALLRILYLDRO 745
611 ---SAQGPAPKRRSQI-LPASGRSLSLPS-----QFSSTASVLPGLSLTTIAP 656
746 DPSDALSKLQOVE-GVNDK---MSOCTSTIPSSSEKQPKIKTEISEGSGDLNL 800
657 TPQD--DSOCCQSPDFGCHDRQLRLLSQPIQPMMPESCDARP--SEVSRTGROYKAYQ 711
801 DALIGDITSSDFYNNSSISSNGSHLGTKQY---FOGTNSLKGSSQVOSTIRPYNRAV 856
712 SQVMPSPDHPHTNSASASRPVLLMG--QAVLHPSFPASRPSPLOQPAQAOQO-PPRY----- 764
857 SLDSFVSVGVSSPP 869
765 -LQAPFTSLHSEQ 776

RESULT 15
US-08-885-291-53
Sequence 53, Application US/08885291A
Patent No. 6057125

GENERAL INFORMATION:
APPLICANT: Takahashi, Joseph S.
APPLICANT: Turek, Fred W.
APPLICANT: Pinto, Lawrence H.
TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
FILE REFERENCE: 0290-5
CURRENT APPLICATION NUMBER: US/08/885,291A
CURRENT FILING DATE: 1997-06-30
EARLIER APPLICATION NUMBER: 08/816,693
EARLIER FILING DATE: 1997-03-13
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 53
LENGTH: 816
TYPE: PRT
ORGANISM: Mus musculus
US-08-885-291-53

Query Match 3.4%; Score 254.5; DB 3; Length 816;

Best Local Similarity 20.8%; Pred. No. 1.8e-10; Matches 190; Conservative 140; Mismatches 352; Indels 231; Gaps 45;

34 EKRRROESKYIEELAEELISANLSDIDNENVPKDKALKEVROIROKEGKTIISND 93
18 EKRRROQEVVLKELSSMLPGNIR-----KMDKITVLEKVIIGLQKINE---VSAQT 66
94 DVQKADVSSTGO-GVIDKDSLGPILLQALDGLFVYNREANIVFVSENVTOYLQKQEDL 152
67 EI--CDIQQDMKPSFLSNEEFTQMLLEALDGEVIYVTTGDSITIVSDSTPLGLHPADV 124
153 VMTSVYNIHEEDRKDFLNLKSTVNGVSWTNEPQROKSHF--FNCMLMKTPTDILE 209
125 MDONLNLFLPEQHSFVYKILSSHML--VTDSPPEFLKSDNDEFLYCHLLNG----- 175
210 DINASPEMRQREYTMQCF-----ALSOPRAMEEGEDLOSCMICA 250
176 --SLNPKPEPTYEYIKFVGNFRSYNNVPSPCNGDNTLSRP-CHVPLGKV--CFIATV 230
251 RRIIT--TGERTFPSPNP-ESFTRHDLGKVVNIDTNS-----LRSSMRGFE- 294
231 RLATPOFLKEMCVADEPLEEFTSRHSLEWKFLDHRAPPIIGYLPFEVLGTS--GYNY 287
295 -----DIRRCIORFFSLNDGOS-----WSQKHGYQAY--LNGHAETPY 333
288 YIIDDELLARCHOHLMOFGKSCCYRFLTKGQWIMLQ-THYITTYHOMNSKPEFTVC 346
334 RFLSLADGITVIAQTQSKL-FRNPVTNDRHGFVSTHFLQREONGYRPNPNVGO--GIRP 389
347 THSVSYADAVRERROELALEDDPPEAMH-----PSAVKEXKSSLEP 388
390 PMAGCNSSVGVGSMSPNOGLQMPSSRAYGLADPSTTGOMSGARYGSSNINASTLPQGMQ 449
389 P-----QPFNALDM-----GASGLPS-SPSPSAS 411
450 SPSSYONNNYGLNMSPPHSGPLAPNOONIMISPRNGSPKIASHQSPVAVGSPMAS 509
412 SRSSHKSSHTAM--SEPTSTPTKLAENSTALPRATLP-----QELPVQGL--SOAA 461
510 SCNTGNHSESSSSLS--ALQAISEGVGTSLLSTLSSPGKLDNSPNNITOPSKVSNOD 566
462 TMTPTALHSSASCDLTQKLLQSLPQ--TGL--QSPAPVYQFSAQFSMOTIK----- 510
567 SKSPGLGYCDQNPVSSMQOSNRDHLSDKESKESVGEAENQGRPLESKGHKKLLQLT 626
511 -----DQLEQRTRIQANIR-----WQBELKHIQEOCLCYODSNVQMFLOQPA 554
627 CSSDDRGHSLNPLSDSCKESSSVTSPS-GVSSSTSGVSTSNMHGSLDEKXRL 685

```

Db 555 VS---LSFSSIQRPAAQOOIQRPAAQOPOLVNTPILOGOITSTOVNQHILRESNVI- 610
QY 686 HKLLQNGNSPAEYAKITAGATGKDTSSITSCGNGNVYKOEQLSPKKENNALLRYLDRD 745
Db 611 ---SAOGPKPMRSSOL-LPASGRSLSLPS-----QFSSTASVLPPLGLSTFTIAP 656
QY 746 DPSDALSKELQOVE-GVDNK---MSQCTSTIPSSQEKDPKIKETSEEGSGDIDLNL 800
Db 657 TPQD--DSQOQPSPDFGHDQLRLLSQPIQPMMPGSCDAROP---SEVSRIGQYKYYAQ 711
QY 801 DAILGDLTSSDFYNNSSISNGSHLGTQOY---FOGTNSLGLKSSQSVOSIRPPYNNRAV 856
Db 712 SQVMFPSPDSHPNTSSASTFVLLMG--QAVLHPSFPASRPSPLQPAQAOQO-PPPY----- 764
QY 857 SLDSPYSVGSSPP 869
Db 765 -LQAPTSLHSEOP 776

```

Search completed: April 19, 2001, 16:50:48
 Job time: 296 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 16:45:53 ; Search time 147.68 Seconds
(without alignments)
1541.232 Million cell updates/sec

Title: US-09-041-994-2

Perfect score: 7383

Sequence: 1 MSGIGENDPLASDSRRKRL.....MNMNPMPSGMPKGPDKKXC 1415

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7383	100.0	1415	US-09-041-994-2	Sequence 2, Appl1
2	7375	99.6	1415	US-09-513-066-2	Sequence 2, Appl1
3	7352.5	99.6	1420	US-09-125-635-4	Sequence 4, Appl1
4	6077.5	82.3	1402	US-09-125-635-12	Sequence 12, Appl1
5	5864.5	79.4	1398	US-09-513-066-14	Sequence 14, Appl1
6	5411	73.3	1402	US-09-445-353A-2	Sequence 2, Appl1
7	5041	68.3	1391	US-09-513-066-15	Sequence 8, Appl1
8	4938	66.3	951	US-09-125-635-8	Sequence 15, Appl1
9	2740.5	37.1	1464	US-08-891-640-2	Sequence 2, Appl1
10	2717.5	36.8	1462	US-09-445-353A-3	Sequence 3, Appl1

11	1966.5	26.6	1127	23	US-60-258-273-113	Sequence 113, App
12	957	13.0	186	15	US-09-125-635-2	Sequence 2, Appl1
13	873	11.8	1036	12	US-08-891-640-3	Sequence 3, Appl1
14	872	11.8	1061	11	US-08-701-154A-5	Sequence 5, Appl1
15	742	10.1	149	23	US-60-236-359-19468	Sequence 19468, A
16	739.5	10.0	147	11	US-08-766-263-2	Sequence 2, Appl1
17	739.5	10.0	147	13	US-08-766-263A-2	Sequence 2, Appl1
18	739.5	10.0	147	13	US-08-988-232-2	Sequence 2, Appl1
19	739.5	10.0	147	17	US-09-398-828-134	Sequence 134, App
20	644	8.7	124	23	US-60-182-515-11	Sequence 11, Appl
21	636.5	8.6	125	23	US-60-177-667-98	Sequence 98, Appl
22	530.5	7.2	106	23	US-60-196-714-414	Sequence 414, App
23	396.5	5.4	1778	23	US-60-191-637-7056	Sequence 7056, App
24	396.5	5.4	1778	23	US-60-191-681-5523	Sequence 5523, App
25	379.5	5.1	1625	23	US-60-173-464-5755	Sequence 5755, App
26	376	5.1	73	15	US-09-125-635-3	Sequence 3, Appl1
27	351.5	4.8	3275	23	US-60-191-637-37720	Sequence 37720, A
28	351.5	4.8	3275	23	US-60-191-637-29308	Sequence 29308, A
29	347.5	4.7	4007	23	US-60-173-464-19825	Sequence 19825, A
30	337.5	4.6	75	23	US-60-196-714-529	Sequence 529, App
31	334	4.5	2976	23	US-60-171-627-1856	Sequence 1856, App
32	334	4.5	2976	23	US-60-173-464-26918	Sequence 26918, A
33	334	4.5	2977	23	US-60-191-637-34837	Sequence 34837, A
34	334	4.5	2977	23	US-60-191-681-27367	Sequence 27367, A
35	333	4.5	5222	23	US-60-245-201-208	Sequence 208, App
36	329.5	4.5	3059	23	US-60-167-217-21631	Sequence 21631, A
37	329.5	4.5	3059	23	US-60-171-627-1719	Sequence 1719, App
38	329.5	4.5	3059	23	US-60-173-464-17597	Sequence 17597, A
39	325.5	4.4	1366	23	US-60-191-637-27784	Sequence 27784, A
40	325.5	4.4	1366	23	US-60-191-681-22393	Sequence 22393, A
41	325.5	4.4	1594	23	US-60-191-637-17909	Sequence 17909, A
42	325.5	4.4	1594	23	US-60-191-681-11162	Sequence 11162, A
43	325	4.4	2280	20	US-09-619-684-624	Sequence 624, App
44	325	4.4	2280	23	US-60-167-217-11753	Sequence 11753, App
45	325	4.4	2280	23	US-60-171-627-975	Sequence 975, App

ALIGNMENTS

RESULT 1
US-09-041-994-2
Sequence 2, Application US/09041994
GENERAL INFORMATION:
APPLICANT: Chen, J. Don
TITLE OF INVENTION: Transcriptional Coactivator for Nuclear
TITLE OF INVENTION: Hormone Receptors
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive and Cockfield
STREET: 28 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,994
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Liepmann, W. Hugo
REGISTRATION NUMBER: 20,407
REFERENCE/DOCKET NUMBER: UMM-026-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-227-7400
TELEFAX: 617-742-4214
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
; LENGTH: 1415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-994-2

Query Match 100.0%; Score 7383; DB 14; Length 1415;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGLGENDPLASDSRKRRLPCDTPGQGLTCSGKERRRQESKYIEBELAELISANLSDID 60
DB 1 MSGLGENDPLASDSRKRRLPCDTPGQGLTCSGKERRRQESKYIEBELAELISANLSDID 60
QY 61 NFNVKPKCALIKETVROIRQIKEGKTISNDDVQKADVSTGGGVIDKDSLGPLLLQA 120
DB 61 NFNVKPKCALIKETVROIRQIKEGKTISNDDVQKADVSTGGGVIDKDSLGPLLLQA 120
QY 121 LDGFLFVNREANIVFSENVTQYLOYKOEDLVNTSVYNIILHEEDRKDFLKNLPRKSTVNG 180
DB 121 LDGFLFVNREANIVFSENVTQYLOYKOEDLVNTSVYNIILHEEDRKDFLKNLPRKSTVNG 180
QY 181 VMTNPPOROKSHTEPCRLMKTPHDILEDINASPEMRGRYETMQCFALSOGRAMMEEE 240
DB 181 VMTNPPOROKSHTEPCRLMKTPHDILEDINASPEMRGRYETMQCFALSOGRAMMEEE 240
QY 241 DLQSCMICARITTEGERTFPSPNESFITRHDLSGKVVNIIDTNSLRSMRPFEDIIRRC 300
DB 241 DLQSCMICARITTEGERTFPSPNESFITRHDLSGKVVNIIDTNSLRSMRPFEDIIRRC 300
QY 301 IORFESLNDGQSMQKRHYOEAYLNGHAETPVYRESLADGTVTAQTKSKLEFRNPVTNDR 360
DB 301 IORFESLNDGQSMQKRHYOEAYLNGHAETPVYRESLADGTVTAQTKSKLEFRNPVTNDR 360
QY 361 HGFVSTHFLQRONGRPNPNVGGGIRPPMAGCNSSVGMGMSPMQGLQMSRRAYGLA 420
DB 361 HGFVSTHFLQRONGRPNPNVGGGIRPPMAGCNSSVGMGMSPMQGLQMSRRAYGLA 420
QY 421 DPTTGQMSGARYGSSNLTSLPGMGQSPSSYQNNNTGLNMSPPHSGPLAQNQNI 480
DB 421 DPTTGQMSGARYGSSNLTSLPGMGQSPSSYQNNNTGLNMSPPHSGPLAQNQNI 480
QY 481 MISPRNRGSPKTAHQFSPVAGVHSPMSSGNTGNHSESSSSLSALQAISEGVGTSILST 540
DB 481 MISPRNRGSPKTAHQFSPVAGVHSPMSSGNTGNHSESSSSLSALQAISEGVGTSILST 540
QY 541 LSSPGKLDNSPNMNTOPSKVSNODSKSPGLFYCDONFVSSMCOGNSRDHLSDEKSE 600
DB 541 LSSPGKLDNSPNMNTOPSKVSNODSKSPGLFYCDONFVSSMCOGNSRDHLSDEKSE 600
QY 601 SSVGEAENGRPLESGKHKLLQTLTCSDDRGHSLTNSPLDSSCKESSVYTPSPGVS 660
DB 601 SSVGEAENGRPLESGKHKLLQTLTCSDDRGHSLTNSPLDSSCKESSVYTPSPGVS 660
QY 661 SSTSGVSTSMHSGSLQEKHRIILHKLQNGNSPAEVAKITAQATGKDTSSITSCGDGN 720
DB 661 SSTSGVSTSMHSGSLQEKHRIILHKLQNGNSPAEVAKITAQATGKDTSSITSCGDGN 720
QY 721 VKKQOLSPKKKENNALLPYLLDRDDPSALSKELQPOVEGYDNKMSQCTSSITPSSOE 780
DB 721 VKKQOLSPKKKENNALLPYLLDRDDPSALSKELQPOVEGYDNKMSQCTSSITPSSOE 780
QY 781 KDPKIKTETSEEGSGDLNDAIILGDLTSSDFYNNSSISNGSHLGTQOVPFGTNSLGK 840
DB 781 KDPKIKTETSEEGSGDLNDAIILGDLTSSDFYNNSSISNGSHLGTQOVPFGTNSLGK 840
QY 841 SSQSVQSTIRPPYRAVSLDSPVSVSSPPVKNIISAFPMMLPKQPMLGPNRMDSQENYGS 900
DB 841 SSQSVQSTIRPPYRAVSLDSPVSVSSPPVKNIISAFPMMLPKQPMLGPNRMDSQENYGS 900
QY 901 SMGGRNNTVTVQTPTSSGDMGLPNSKAGMEPMNSMGRPGGDVNTSLPRALGSIPT 960

DB 901 SMGGRNNTVTVQTPTSSGDMGLPNSKAGMEPMNSMGRPGGDVNTSLPRALGSIPT 960
QY 961 LPLRNSISGARPVLOQOQOMLQMRGELIPMGGANPYGOAASNOLGSPMGMLSMEOY 1020
DB 961 LPLRNSISGARPVLOQOQOMLQMRGELIPMGGANPYGOAASNOLGSPMGMLSMEOY 1020
QY 1021 SHGTQNRPLLRNSLDDLVGPPSNLEGQSDERALLDQHLTSLNTDATGTEEDRALGIPE 1080
DB 1021 SHGTQNRPLLRNSLDDLVGPPSNLEGQSDERALLDQHLTSLNTDATGTEEDRALGIPE 1080
QY 1081 LVNQGALERPQKDAFQGGAAVMMQKAGLYGQTYPAQGPMPQGFHLOQGSFSPNSMKN 1140
DB 1081 LVNQGALERPQKDAFQGGAAVMMQKAGLYGQTYPAQGPMPQGFHLOQGSFSPNSMKN 1140
QY 1141 QMNQGNFPLQGMHPRANIMRPTNTPKQLRMQLQRLQGOQFLMQSRALELKNENPTA 1200
DB 1141 QMNQGNFPLQGMHPRANIMRPTNTPKQLRMQLQRLQGOQFLMQSRALELKNENPTA 1200
QY 1201 GGAAYVRPMQPOQGFLLNQVYAQNSRELLSHHFQOQRYVAMM00000000000000 1260
DB 1201 GGAAYVRPMQPOQGFLLNQVYAQNSRELLSHHFQOQRYVAMM00000000000000 1260
QY 1261 Q00000Q0TAFSPRPVVTASPMQGLAGPTMPOAPQOFPYQPVYMGQCPDPAFGRVS 1320
DB 1261 Q00000Q0TAFSPRPVVTASPMQGLAGPTMPOAPQOFPYQPVYMGQCPDPAFGRVS 1320
QY 1321 SPPNMMSSRMQSPQNPMMQHPOAASIVQSSSEMKMPSGNLARNSFSQOQFAHOGNPAY 1380
DB 1321 SPPNMMSSRMQSPQNPMMQHPOAASIVQSSSEMKMPSGNLARNSFSQOQFAHOGNPAY 1380
QY 1381 YSMVHNGSSGHMGQNNMNPMPGMPGPDQKTC 1415
DB 1381 YSMVHNGSSGHMGQNNMNPMPGMPGPDQKTC 1415

RESULT 2
US-09-513-066-2
; Sequence 2, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Leo, Christopher
; APPLICANT: Li, Hui
; TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
; FILE REFERENCE: UMG-026CP
; CURRENT APPLICATION NUMBER: US/09/513, 066
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: USSN 09/041, 994
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: USSN 60/073, 674
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 1415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-066-2

Query Match 99.9%; Score 7375; DB 19; Length 1415;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSGLGENDPLASDSRKRRLPCDTPGQGLTCSGKERRRQESKYIEBELAELISANLSDID 60
DB 1 MSGLGENDPLASDSRKRRLPCDTPGQGLTCSGKERRRQESKYIEBELAELISANLSDID 60
QY 61 NFNVKPKCALIKETVROIRQIKEGKTISNDDVQKADVSTGGGVIDKDSLGPLLLQA 120
DB 61 NFNVKPKCALIKETVROIRQIKEGKTISNDDVQKADVSTGGGVIDKDSLGPLLLQA 120


```
QY 121 LDGFLFVNEANIVFSENVTOYLOKQEDLVNTSVYNIHEDRDKFLKNLPKSTVNG 180
DB 121 LDGFLFVNEANIVFSENVTOYLOKQEDLVNTSVYNIHEDRDKFLKNLPKSTVNG 180
QY 181 VSWTNEPQROKSHTFNCRLMLKTPHDLIEDINASPEMRORYETMOCFALSOPRAMEEGE 240
DB 181 VSWTNEPQROKSHTFNCRLMLKTPHDLIEDINASPEMRORYETMOCFALSOPRAMEEGE 240
QY 241 DLOSCMCVARRITTTGERTPPSPNESEFTTRHDLISGKVVNIDTNSLRSSMRPGEFEDIIIRC 300
DB 241 DLOSCMCVARRITTTGERTPPSPNESEFTTRHDLISGKVVNIDTNSLRSSMRPGEFEDIIIRC 300
QY 301 IOFFESLNDQOSMSOKRHVQEAFLNGHAETPVYREFSLADGTIYTAQTKSLFRNPVTNDR 360
DB 301 IOFFESLNDQOSMSOKRHVQEAFLNGHAETPVYREFSLADGTIYTAQTKSLFRNPVTNDR 360
QY 361 HGFVSTHFLQREONGYRPNPNVGOGIRPPMACNSSVSGMSMSPNOGLQMPSSRAYGLA 420
DB 361 HGFVSTHFLQREONGYRPNPNVGOGIRPPMACNSSVSGMSMSPNOGLQMPSSRAYGLA 420
QY 421 DPSTTGMGSGARYGSSNIALSLTPPGMOSPSSYQNNNTYGLNASSPPHSGPGLAPNOQNT 480
DB 421 DPSTTGMGSGARYGSSNIALSLTPPGMOSPSSYQNNNTYGLNASSPPHSGPGLAPNOQNT 480
QY 481 MISPRNNGSPKIASHOFSPVAGVHSPMASSGNTGNHSFSSSSLSALQAISEGVGTSLST 540
DB 481 MISPRNNGSPKIASHOFSPVAGVHSPMASSGNTGNHSFSSSSLSALQAISEGVGTSLST 540
QY 541 LSSPBGKLDNSPMNNTTQPSKVSNOQSKSPLGFYCDQNPVESSMCOQNSRBDHLSDEKSE 600
DB 541 LSSPBGKLDNSPMNNTTQPSKVSNOQSKSPLGFYCDQNPVESSMCOQNSRBDHLSDEKSE 600
QY 601 SSVGAEONQGPLESKHKRLQLDITCSSDDRGHSSLTNPDLSSCKESSVYSPBGVS 660
DB 601 SSVGAEONQGPLESKHKRLQLDITCSSDDRGHSSLTNPDLSSCKESSVYSPBGVS 660
QY 661 SSTSGCVSSTSNMHSGLDQEKHRLKHLONGNSPAEVAKITTAQATKQDSTISGCDGN 720
DB 661 SSTSGCVSSTSNMHSGLDQEKHRLKHLONGNSPAEVAKITTAQATKQDSTISGCDGN 720
QY 721 VVKQEOQLSPKKENNALRLYLDRDPSDALSKELQPOVEGVNKMISOCTSTIPSSSOE 780
DB 721 VVKQEOQLSPKKENNALRLYLDRDPSDALSKELQPOVEGVNKMISOCTSTIPSSSOE 780
QY 781 KDKKITEETSEEGSDLDNDAILGDLTSSDFYNNSSNGSHLGTQOYVFOGTNSLGLK 840
DB 781 KDKKITEETSEEGSDLDNDAILGDLTSSDFYNNSSNGSHLGTQOYVFOGTNSLGLK 840
QY 841 SSOSVOSIRPPYRAVSLDSPVSGSSPVKNISAFPMLEKOPMLGPNPMDSOEYGS 900
DB 841 SSOSVOSIRPPYRAVSLDSPVSGSSPVKNISAFPMLEKOPMLGPNPMDSOEYGS 900
QY 901 SMGCPNKNVTVTQTPSSGDMGLPNSKAGRMEMPNSNSMGRPGDDYNTSLRPAAGSIPT 960
DB 901 SMGCPNKNVTVTQTPSSGDMGLPNSKAGRMEMPNSNSMGRPGDDYNTSLRPAAGSIPT 960
QY 961 LPIRBSNIPGARVYLOQOQOQMLQMRBEITPMGCANPYGGAASNOUGSPMDGLSMEY 1020
DB 961 LPIRBSNIPGARVYLOQOQOQMLQMRBEITPMGCANPYGGAASNOUGSPMDGLSMEY 1020
QY 1021 SHGTQNPFLRLNSLDLVGPPSNLEGSDELRALDQHLTLLSTWDATGEIIRALGIPRE 1080
DB 1021 SHGTQNPFLRLNSLDLVGPPSNLEGSDELRALDQHLTLLSTWDATGEIIRALGIPRE 1080
QY 1081 LVNQGQALEKQDAFOGQEAAYMMQKAGLYGOTYPAQGPMPQGGFHLQOSPSFSNMAN 1140
DB 1081 LVNQGQALEKQDAFOGQEAAYMMQKAGLYGOTYPAQGPMPQGGFHLQOSPSFSNMAN 1140
QY 1141 QMNOQGNFPLQGMHPRANIRPRTNTPKOLRMLOLQOQOFLNOSROALETKMEMP7A 1200
DB 1141 QMNOQGNFPLQGMHPRANIRPRTNTPKOLRMLOLQOQOFLNOSROALETKMEMP7A 1200
QY 1201 GGAAYVNRPMQPOQGFLLNAQMVARSRELLSHHFROQRVAMMMQOQOQOQOQOQOQ 1260
DB 1201 GGAAYVNRPMQPOQGFLLNAQMVARSRELLSHHFROQRVAMMMQOQOQOQOQOQOQ 1260

QY 1201 GGAAYVNRPMQPOQGFLLNAQMVARSRELLSHHFROQRVAMMMQOQOQOQOQOQOQ 1260
DB 1201 GGAAYVNRPMQPOQGFLLNAQMVARSRELLSHHFROQRVAMMMQOQOQOQOQOQOQ 1260
QY 1261 QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 1320
DB 1261 QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 1320
QY 1321 SPPNMMSSRMGSPQNPMMQHPQAASITQOSSEKMGWSSGNLARRSSSTQOQFAHQGNAY 1380
DB 1321 SPPNMMSSRMGSPQNPMMQHPQAASITQOSSEKMGWSSGNLARRSSSTQOQFAHQGNAY 1380
QY 1381 YSMVHMGSSGHMGOMNMNPMPSGMPGPDQKXC 1415
DB 1381 YSMVHMGSSGHMGOMNMNPMPSGMPGPDQKXC 1415

RESULT 3
US-09-125-635-4
; Sequence 4, Application US/09125635
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; FILE OF INVENTION: A1B1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; PRIOR APPLICATION NUMBER: 60/049,728
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1420
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-125-635-4

Query Match 99.6%; Score 7352.5; DB 15; Length 1420;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1411; Conservative 2; Mismatches 2; Indels 5; Gaps 1;

QY 1 MSGIGENIDPLASDRKRLPCDTPGQGLTCSGSKRRREDESYIEELALISANLSDID 60
DB 1 MSGIGENIDPLASDRKRLPCDTPGQGLTCSGSKRRREDESYIEELALISANLSDID 60
QY 61 NFNVKPKCAILKETVQRIQIRKQGTISNDDVQKADVSSYGQYVIDKOSLGPLLLQA 120
DB 61 NFNVKPKCAILKETVQRIQIRKQGTISNDDVQKADVSSYGQYVIDKOSLGPLLLQA 120
QY 121 LDGFLFVNEANIVFSENVTOYLOKQEDLVNTSVYNIHEDRDKFLKNLPKSTVNG 180
DB 121 LDGFLFVNEANIVFSENVTOYLOKQEDLVNTSVYNIHEDRDKFLKNLPKSTVNG 180
QY 181 VSWTNEPQROKSHTFNCRLMLKTPHDLIEDINASPEMRORYETMOCFALSOPRAMEEGE 240
DB 181 VSWTNEPQROKSHTFNCRLMLKTPHDLIEDINASPEMRORYETMOCFALSOPRAMEEGE 240
QY 241 DLOSCMCVARRITTTGERTPPSPNESEFTTRHDLISGKVVNIDTNSLRSSMRPGEFEDIIIRC 300
DB 241 DLOSCMCVARRITTTGERTPPSPNESEFTTRHDLISGKVVNIDTNSLRSSMRPGEFEDIIIRC 300
QY 301 IOFFESLNDQOSMSOKRHVQEAFLNGHAETPVYREFSLADGTIYTAQTKSLFRNPVTNDR 360
DB 301 IOFFESLNDQOSMSOKRHVQEAFLNGHAETPVYREFSLADGTIYTAQTKSLFRNPVTNDR 360
QY 361 HGFVSTHFLQREONGYRPNPNVGOGIRPPMACNSSVSGMSMSPNOGLQMPSSRAYGLA 420
DB 361 HGFVSTHFLQREONGYRPNPNVGOGIRPPMACNSSVSGMSMSPNOGLQMPSSRAYGLA 420
QY 421 DPSTTGMGSGARYGSSNIALSLTPPGMOSPSSYQNNNTYGLNASSPPHSGPGLAPNOQNT 480
DB 421 DPSTTGMGSGARYGSSNIALSLTPPGMOSPSSYQNNNTYGLNASSPPHSGPGLAPNOQNT 480
QY 481 MISPRNNGSPKIASHOFSPVAGVHSPMASSGNTGNHSFSSSSLSALQAISEGVGTSLST 540
DB 481 MISPRNNGSPKIASHOFSPVAGVHSPMASSGNTGNHSFSSSSLSALQAISEGVGTSLST 540
```

```
|||||
Db 481 M1SPNRKSPK1ASHQFSVAVGVSMASSGNTGNHSFSSSSLSALQALISEVGTSLST 540
Qy 541 LSSPGKLDNSPNMNTITOPSKVSNODSKSPGLFCYCDQNFVESSMCOQNSRDHLDKESKE 600
Db 541 LSSPGKLDNSPNMNTITQSKVSNODSKSPGLFCYCDQNFVESSMCOQNSRDHLDKESKE 600
Qy 601 SSVGEAENGRGLESKGHKLLQLLTCCSDDRGHSLTSPDLSSCKESSVYTSFSGVS 660
Db 601 SSVGEAENGRPLESKGHKKLLQLLTCCSDDRGHSLTSPDLSSCKESSVYTSFSGVS 660
Qy 661 SSTSGSVSTSMHGSLLQEKHRLHKLLQNGNSPAEVAKITAQATKXTSSTTSCGDN 720
Db 661 SSTSGSVSTSMHGSLLQEKHRLHKLLQNGNSPAEVAKITAQATKXTSSTTSCGDN 720
Qy 721 VVKOEQLSPKKENNALRYLLDRDDPSDALSKELQPOVEGVNKM5QCTSTTIPSSQ 780
Db 721 VVKOEQLSPKKENNALRYLLDRDDPSDALSKELQPOVEGVNKM5QCTSTTIPSSQ 780
Qy 781 KKPRTKTESESGDLDLDAIILGDLTSSDFYNMNSISSNGSHLTKQVFCGTNSLGLK 840
Db 781 KKPRTKTESESGDLDLDAIILGDLTSSDFYNMNSISSNGSHLTKQVFCGTNSLGLK 840
Qy 841 SSSQSVOSTRPPYNAVSLDSPSVSGSSPPVKNISAFPMILPKQPMILGPNRMD5QENYGS 900
Db 841 SSSQSVOSTRPPYNAVSLDSPSVSGSSPPVKNISAFPMILPKQPMILGPNRMD5QENYGS 900
Qy 901 SMGPNRNTVLTQTPSSGDMGLPNSKAGHMEPMNSMSMRPGGDYVTS1LRPALGSIPT 960
Db 901 SMGPNRNTVLTQTPSSGDMGLPNSKAGHMEPMNSMSMRPGGDYVTS1LRPALGSIPT 960
Qy 961 LPLRNSITGARPVLOQOQOMLQMRPGETPMGNGANPYQAAASNOLG5WPMGMLSMEDV 1020
Db 961 LPLRNSITGARPVLOQOQOMLQMRPGETPMGNGANPYQAAASNOLG5WPMGMLSMEDV 1020
Qy 1021 SHGTORPLRLRNSLDDLVGPPSNLEGSQDERALLDQLLHTLSTNTATG1LEIDRALG1PE 1080
Db 1021 SHGTORPLRLRNSLDDLVGPPSNLEGSQDERALLDQLLHTLSTNTATG1LEIDRALG1PE 1080
Qy 1081 LVNQGALPEPKODAFQGOEAAVMMQKAGLYGQTYPAQGRPMQOGFH1LQGSQSPFNSMKN 1140
Db 1081 LVNQGALPEPKODAFQGOEAAVMMQKAGLYGQTYPAQGRPMQOGFH1LQGSQSPFNSMKN 1140
Qy 1141 QMNQOGNFPLOGMHPRANIMRRTNTPKOLRMQLODRLOGOFLNDSRQALELKMENPTA 1200
Db 1141 QMNQOGNFPLOGMHPRANIMRRTNTPKOLRMQLODRLOGOFLNDSRQALELKMENPTA 1200
Qy 1201 GGAAYWRPMQPOQGF1NAQNTAORSRELLSHHFROQRYAMMM-----00000000000 1255
Db 1201 GGAAYWRPMQPOQGF1NAQNTAORSRELLSHHFROQRYAMMM-----00000000000 1255
Qy 1255 QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1315
Db 1255 QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 1315
Qy 1315 FERRVSSPPRMAHMSRKGQPNQPMQHPQAASITYOSSEKMGWSGNLARNSSFSQOQFAHQ 1375
Db 1315 FERRVSSPPRMAHMSRKGQPNQPMQHPQAASITYOSSEKMGWSGNLARNSSFSQOQFAHQ 1375
Qy 1376 GNPAYVSNVHMGSSGHMGOMMNPMPMGPMGPMGPOKTC 1415
Db 1376 GNPAYVSNVHMGSSGHMGOMMNPMPMGPMGPMGPOKTC 1415
Qy 1415 GNPAYVSNVHMGSSGHMGOMMNPMPMGPMGPMGPOKTC 1420
Db 1415 GNPAYVSNVHMGSSGHMGOMMNPMPMGPMGPMGPOKTC 1420
```

```

; PRIOR APPLICATION NUMBER: 60/049,728
; BEST FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-125-635-12

Query Match      82.3%  Score 6077.5  DB 15  Length 1402:
Best Local Similarity 81.7%  Pred No. 0:
Matches 1175; Conservative 85; Mismatches 118; Indels 61; Gaps 11:

Qy 1 MSGLGE-NIDPLASRRKKRLPCDTPPGGLTCSGEKRRRQESKYIEELAEILSANDSDI 59
Db 1 MSGLGESSLDPLAESRRKKRLPCDAPGGLVYSGEKWRREQSKYIEELAEILSANDSDI 60
Qy 60 DNFNPKPDKCALIKETVROIROIKEDGKTI5NDVDYQKADVSTGCGVTDKDSLGLPLLIQ 119
Db 61 DNFNPKPDKCALIKETVROIROIKEDGKTI5DDVDYQKADVSTGCGVTDKDSLGLPLLIQ 120
Qy 120 ALDGLFLVYVNRANIVFVSENTVOYLOVKOEQLVMTSVYNTLHEDRDKDLKNLPRKSTYN 179
Db 121 ALDGLFLVYVNRANIVFVSENTVOYLOVKOEQLVMTSVYNTLHEDRDKDLKHLPRKSTYN 180
Qy 180 GVSMTNEPQROKSHTFNCBMLKTPHDILEDINASPEMRORYETMQCFAL5QPRAMMEBG 239
Db 181 GVSMTNEPQROKSHTFNCBMLKTPHDILEDINASPEMRORYETMQCFAL5QPRAMMEBG 239
Qy 240 EDLQSCMTCVARRITGERTPEPSNPBESFTTRHDL5GKVVNIDTNSLRSSMRPGFEDIIR 299
Db 240 EDLQSCMTCVARRITGERTPEPSNPBESFTTRHDL5GKVVNIDTNSLRSSMRPGFEDIIR 296
Qy 300 C1QRFPSLNDGQSW5QKRYHGOEAYLNGHAEPTVYRSLADGTVIAQTKSLFRBPVTMD 359
Db 297 C1QRFPSLNDGQSW5QKRYHGOEAYLNGHAEPTVYRSLADGTVIAQTKSLFRBPVTMD 356
Qy 360 RHGFVSTHFLQEQNGYRPNPVPVGGIRPMAQCNSSVSGMSMSPNOGLQMPSSRAYGL 419
Db 357 RHGFVSTHFLQEQNGYRPNPVPVGGIRPMAQCNSSVSGMSMSPNOGLQMPSSRAYGL 411
Qy 420 ADPSTTGQSGARYGSSNIA5LTPGQSPSTIONNNTYGLMSSPPRGSGGLAPNOON 479
Db 412 ADPSTTGQSGARYGSSNIA5LTPGQSPSTIONNNTYGLMSSPPRGSGGLAPNOON 471
Qy 480 IMISPRNRSKPRTASHQFSVAVGVSMASSGNTGNHSFSSSSLSALQALISEVGTSLIS 539
Db 472 IMISPRNRSKPRTASHQFSVAVGVSMASSGNTGNHSFSSSSLSALQALISEVGTSLIS 531
Qy 540 T1LSPGPKLDNSPNMNTITOPSKVSNODSKSPGLFCYCDQNFVESSMCOQNSRDHLDKESK 599
Db 532 T1LSPGPKLDNSPNMNTISQPSKVSGODSKSPGLFCYCDQNFVESSMCOQNSRDHLDKESK 591
Qy 600 ESSVEGAENGRPLESKGHKKLLQLLTCCSDDRGHSLTSPDLSSCKESSVYTSFSGV 659
Db 592 ESSVEGAENGRPLESKGHKKLLQLLTCCSDDRGHSLTSPDLSSCKESSVYTSFSGV 651
Qy 660 SSTSGSVSTSMHGSLLQEKHRLHKLLQNGNSPAEVAKITAQATKXTSSTTSCGG 719
Db 652 SSTSGSVSTSMHGSLLQEKHRLHKLLQNGNSPAEVAKITAQATKXTSSTTSCGG 711
Qy 720 NVVKOEQLSPKKENNALRYLLDRDDPSDALSKELQPOVEGVNKM5QCTSTTIPSSQ 779
Db 712 NVVKOEQLSPKKENNALRYLLDRDDPSDALSKELQPOVEGVNKM5QCTSTTIPSSQ 770
Qy 780 KKPRTKTESESGDLDLDAIILGDLTSSDFYNMNSISSNGSHLTKQVFCGTNSLGL 839
Db 771 KKPRTKTESESGDLDLDAIILGDLTSSDFYNMNSISSNGSHLTKQVFCGTNSLGL 828
Qy 840 SSSQSVOSTRPPYNAVSLDSPSVSGSSPPVKNISAFPMILPKQPMILGPNRMD5QENY 899
Db 840 SSSQSVOSTRPPYNAVSLDSPSVSGSSPPVKNISAFPMILPKQPMILGPNRMD5QENY 899
```

```

Db 829 RSPQVQSVAPRYNRAVSLDSPVSGPVPKNVSAFPGILPKOPILLAGNRMMDSOENY 888
QY 900 SSMGPNRNTVYTPSGSDMGLPNSKAGMEBMSNSMGPGGDVNTSLPRPALGSI 959
Db 889 ANN-GPRNVPVPTSSPGDGLANSRASMELASSPLRTGADYSATLPRPAMGQSV 947
QY 960 TPLRNSNIPGARVY-----00000MLOMRGEIPMGAN 996
Db 948 TPLRNSNRLPGARSLQ000000000000000000000000MLOMRGELPMGCVN 1007
QY 997 PYGQAASNOUGSPMGMSMEQVSHGTQNRPLLRNSLDDLVPSPNSLEGOSDERALLDQ 1056
Db 1008 PVPAPVSNPOGSPMGMSMEQVSHGTQNRPLLRNSLDDLVPSPNSLEGOSDERALLDQ 1067
QY 1057 LHTLSTNTATGTEIDRALGIPELVNOGALPBPDAPOGGAAYMMODKAGLYGOTY 1116
Db 1068 LHTLSTNTATGTEIDRALGIPELVNOGALPBPDAPOGGAAYMMODKAGLYGOTY 1127
QY 1117 ACPGPMOGEFHLQOSPFSPNSMNMNOQGNFLOGMHPRANIMRPTNTPKOLRMOLQ 1176
Db 1128 ACPGPMOGEFHLQOSPFSPNSMNMNOQGNFLOGMHPRANIMRPTNTPKOLRMOLQ 1187
QY 1177 RLOGOQFLNOSROALEKMEKMPNAGCAAYRPMQPOQGFLLANOVAQSRRELSHHFRQ 1236
Db 1188 RLOGOQFLNOSROALEKMEKMPNAGCAAYRPMQPOQGFLLANOVAQSRRELSHHFRQ 1243
QY 1237 QRYAMMMQ0000000000000000000000000000000000000000000000000 1296
Db 1244 QRYAMMMQ0000000000000000000000000000000000000000000000000 1283
QY 1297 POQFPQPNYGMQOPDPARFGRVSPPNAMSSRMGSPQPMQHOAASITQSSSEKGM 1356
Db 1284 POQFPQPNYGMQOPDPARFGRVSPPNAMSSRMGSPQPMQHOAASITQSSSEKGM 1343
QY 1357 PSGNLARNSFSQOQPAHOGNPAVYSHVHNGSSGGMNMPMPNSGMPGPDQKXC 1415
Db 1344 PSGNLARNSFSQOQPAHOGNPAVYSHVHNGSSGGMNMPMPNSGMPGPDQKXC 1402

RESULT 5
US-09-513-066-14
; Sequence 14, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Leo, Christopher
; APPLICANT: Li, Hui
; TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
; FILE OF INVENTION: STEROID NUCLEAR RECEPTORS
; FILE REFERENCE: UM-026CP
; CURRENT APPLICATION NUMBER: US/09/513,066
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: USSN 09/041,994
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: USSN 60/073,674
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-066-14

```

```

Query Match 79.4%; Score 5864.5; DB 19; Length 1398;
Best Local Similarity 79.7%; Pred. No. 0;
Matches 1145; Conservative 90; Mismatches 141; Indels 61; Gaps 13;

```

```

QY 1 MSGLGE NLDPLASDSKRLPCDTPGCGTSCGEKRRROESKYETIELAEILASISLSDI 59
Db 1 MSGLGESDPLPLAESKRLPCDAPGGLVYSGEKWRROESKYETIELAEILASISLSDI 60
QY 60 DNFNVKPKCALIKETVROJKEGKGTISDDDDVQKADVSTGQGVIDKDSLGPLLQ 119

```

```

Db 61 DNFNVKPKCALIKETVROJKEGKGTISDDDDVQKADVSTGQGVIDKDSLGPLLQ 120
QY 120 ALDGFLEVVNRKANIYVSNVTOYLOKQEDLVNTSVYIILHEDRKDL---KNLPKS 176
Db 121 ALDGFLEVVNRKANIYVSNVTOYLOKQEDLVNTSVYIILHEDRKDL---KNLPKS 180
QY 177 TVNGVSWTNEPOROKSHTFCRMIMKTPHDILIDINASPMPORVETMOCFALSOPRAM 236
Db 181 EFLGLMTRD--KAPILIVRLMKT-HDILEDVNASPETRORYETMOCFALSOPRAM 237
QY 237 EGEDDLOSCMIVARRITTEERTPSPNESEFTRHDLGKRVNIDTNSLRSSMRPGEDI 296
Db 238 EGEDDLOSCMIVARRITTEERTPSPNESEFTRHDLGKRVNIDTNSLRSSMRPGEDI 294
QY 297 IRRICQFFESLNGOSMSQKRYHQAELNHAETPVYRESLAGTYTATQTKKLRNPV 356
Db 295 IRRICQFFESLNGOSMSQKRYHQAELNHAETPVYRESLAGTYTATQTKKLRNPV 354
QY 357 TNDRHGFTSTHFLQREONGYRPNPNVYOGIRPPMAGCNSSVGMSSMPNOGOMPSSRA 416
Db 355 TNDRHGFTSTHFLQREONGYRPNPNVYOGIRPPMAGCNSSVGMSSMPNOGOMPSSRA 409
QY 417 YGLADPSTTGOMSGARYGSSNTIASLTGPGMOSPPSSYQNNNYGLNNSPPHSGPLAPN 476
Db 410 YGVDPDPSNTGOMSGARYGSSNTIASLTGPGMOSPPSSYQNNNYGLNNSPPHSGPLAPN 469
QY 477 QONIMISPRNRGSPKASHOPSPVAGYHSPMASSGNTGHSFSSSLALQAISEGVTS 536
Db 470 QONIMISPRNRGSPKASHOPSPVAGYHSPMASSGNTGHSFSSSLALQAISEGVTS 529
QY 537 LSTLSPGPKLNSPMPNITTOPSKVNSODSKSPILGYCQONVYESSMCOGNSRDHLSK 596
Db 530 LSTLSPGPKLNSPMPNITTOPSKVNSODSKSPILGYCQONVYESSMCOGNSRDHLSK 589
QY 597 ESKESSEVEGAENQGPLESKHKLLQTLCCSSDPRHSSSLTSPPLDSSCKESSVSATSP 656
Db 590 ESKESSEVEGAENQGPLESKHKLLQTLCCSSDPRHSSSLTSPPLDSSCKESSVSATSP 649
QY 657 SGVSSSTSGVSSSTSNMHSGLDQKRLHLKLLONGSNPAEVAKTRQATGKTSTTSC 716
Db 650 SGVSSSTSGVSSSTSNMHSGLDQKRLHLKLLONGSNPAEVAKTRQATGKTSTTSC 709
QY 717 GDGNVYKQEOQLSPKRRKNNMLRLTLDRDPDALSKELOPOVEGYVNNKSOCTSTIPS 776
Db 710 GEG-TTRQEOQLSPKRRKNNMLRLTLDRDPDALSKELOPOVEGYVNNKSOCTSTIPS 768
QY 777 SSOEKDPKIKTETSEEGSDLDNDAITLGLTSSDFYNN--PENGHPGAKQOMFAGPSS 836
Db 769 SSOEKDPKIKTETSEEGSDLDNDAITLGLTSSDFYNN--PENGHPGAKQOMFAGPSS 826
QY 837 LGLKSSQVOSIRPYNRAVSLDSPVSGSPPYKKNISAFPMILPKOMLGNPRMDSOE 896
Db 827 LGLKSSQVOSIRPYNRAVSLDSPVSGSPPYKKNISAFPMILPKOMLGNPRMDSOE 886
QY 897 NYGSSMGPNRNTVYTPSGSDMGLPNSKAGMEBMSNSMGPGGDVNTSLPRPALG 956
Db 887 NYGSSMGPNRNTVYTPSGSDMGLPNSKAGMEBMSNSMGPGGDVNTSLPRPALG 945
QY 957 SIPTPLRNSNIPGARVY-----00000MLOMRGEIPMGANPY 998
Db 946 SIPTPLRNSNIPGARVY-----00000MLOMRGEIPMGANPY 1005
QY 999 GQAASNOUGSPMGMSMEQVSHGTQNRPLLRNSLDDLVPSPNSLEGOSDERALLDQ 1058
Db 1006 SPVQSNPOGSPMGMSMEQVSHGTQNRPLLRNSLDDLVPSPNSLEGOSDERALLDQ 1065
QY 1059 TLLSNTDPTATGTEIDRALGIPELVNOGALPBPDAPOGGAAYMMODKAGLYGOTYPAQ 1118
Db 1066 TLLSNTDPTATGTEIDRALGIPELVNOGALPBPDAPOGGAAYMMODKAGLYGOTYPAQ 1125
QY 1119 GPPMOGFLHOGOSPFSPNSMNMNOQGNFLOGMHPRANIMRPTNTPKOLRMOLQORL 1178

```


PRIOR APPLICATION NUMBER: USSN 09/041,994
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: USSN 60/073,674
PRIOR FILING DATE: 1998-02-04
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 15
LENGTH: 1391
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-066-15

Query Match 68.3%; Score 5041; DB 19; Length 1391;
Best Local Similarity 69.8%; Pred. No. 0;
Matches 998; Conservative 161; Mismatches 217; Indels 54; Gaps 24;

1 MSGIGEN-LDPLASDSRRK-LPCDTPGQGLTSCGKRRRDESKYIEELAEILSANLSD 58
1 MSGIGENSLDPLASERKRRKPPGGLTSCGKRRRDESKYIEELAEILSANLSD 60
59 IDNFNPKPKCALKEVROIRQIKQKTIISDDVQKADSVSTGQVYIDKSLGPLLL 118
61 IDNFNPKPKCALKEVROIRQIKQKTIISDDVQKADSVSTGQVYIDKSLGPLLL 120
119 QALDGLFVYNREANFVSENYTOYLOKQEDLVNTSVYNIHEEDRKDFLNLEKSTY 178
121 QALDGLFVYNREANFVSENYTOYLOKQEDLVNTSVYNIHEEDRKDFLNLEKSTY 180
179 NGVSTNEPOKRSHTFNCRLMKTPIHDLINASPEMROREYEQCFALSOPTAMEE 238
181 NGVSTNEPOKRSHTFNCRLMKTPIHDLINASPEMROREYEQCFALSOPTAMEE 239
239 GEDLOSCMIVARRITTEGERTPEPSPESFTTRHDSGKVVYNIIDTSLRSSMRPFEDTIR 298
240 GEDLOSCMIVARRITTEGERTPEPSPESFTTRHDSGKVVYNIIDTSLRSSMRPFEDTIR 299
299 RCIQRFSLNDGOSMSOKRYOAYLNGAEETPYRPSLADGTIVTAQRTSKLFRNPVTN 358
300 RCIQRFSLNDGOSMSOKRYOAYLNGAEETPYRPSLADGTIVTAQRTSKLFRNPVTN 359
359 DRHGFVSTHFLQEQNGYRPNPVPVGGIRPMAAGCNSSV-GGMSMSPNOGLQMPSSRAY 417
360 DRHGFVSTHFLQEQNGYRPNPVPVGGIRPMAAGCNSSV-GGMSMSPNOGLQMPSSRAY 415
418 GLADPSTQMGARYGSSNTASLTTPGQMSPSSTYQNN-NYGLMSSPFGHSPGLIAPN 476
416 GMDPNSMAQOMQMRKSPGNMAYVQAGVQO-SFYQNNNSNGLMNSPFGHSPGNMAN 474
477 QONIMTSPNRSKPTASHOFSVAGVHSPMASSGNTGNHSPSSSLSLALQISEGVTS 536
475 QPILAVSPNRRASPKMASQFSPVPMNSPMSGSSGSSSLALHAISEGVTS 534
537 LLSLSSPPKPLDNPNNMTTOPSKVSNODSKPPLGFYCDONVESSMCOSSNRDLSDK 596
535 LLSLSSPPKPLDNPNNMTTOPSKVSNODSKPPLGFYCDONVESSMCOSSNRDLSDK 594
597 ESKSEVEGAENORGLPEKHKLLQLLTCCSDDRGHSLSLNSPLDSSCKESSVSTSP 656
595 DYKENIFEESESQSQAESKHKLLQLLTCCSDDRGHSLSLNSPLDSSCKESSVSTSP 651
657 SGVSSSTSGVSTSMHMSLLOEKHRIILKLLQNGNSPRAEVAKITAAQATGKP-----TSS 712
652 SGVSSSTSGVSTSMHMSLLOEKHRIILKLLQNGNSPRAEVAKITAAQATGKP-----TSS 711
713 ITSCDGNVAVKQESLPPKKNALLRYLLDRDPDSDALSKELQPOVEGVDNMSQCTSS 772
712 SAPCTEA-TVKREQLSPKKKNALLRYLLDRDPDSDALSKELQPOVEGVDNMSQCTSS 770
773 TTPSSQEDPKITKTESEBGSDLDNLDAIIGDLTSSDFYNNISISNGSLGCTKQVFO 832
771 NPTTSQEDPKITKTESEBGSDLDNLDAIIGDLTSSDFYNNISISNGSLGCTKQVFO 830

833 GTNSLGLKSSQSVOSIRPPYNNRAVSLDPSVSGSPPVKNISAFMLPKPOMLGNPRMM 892
831 DSFTLAMRSPDSMQSRPPENRAMSLDSD---STTPYRVNVSFPMPLPKQMI-GSPRRM 886
893 DSGENTGSSMG-GPNRNVYVOTQTPSSGDMGLPNSKAGRMENPNNSMGRPGDYNTSLPR 951
887 DGODNFGVMWGSQPNR-SMNQHP-GGDMAQNSAVNRLPEPVGVSVGRGPDYSAMTR 943
952 PALGSIPTPLPLANSIPGARPYLQOQOQMLQRPSEIIPKMKANRYGQAASNLQSGNP 1011
944 PAMGNMPPGLTTRNSNIPSPRPVMOOQOHLPMRPDMAMSNIPYGOAPNPPGSSWP 1003
1012 DGLMSMEVYSHGTONRPLRLNSLDDLVGPSPNSEGSDERALLDOLHTLLSNFDATGLE 1071
1004 DATM-MNQRGGAQNLQGRNSLDDLCPPSTYEGOTDELALDQLHTLLSNFDATGLE 1062
1072 IDRALGIPELVNOGALPEKQDAFOGEAVMMQDQAGLYGOTPYPAQGPMPQ-GGF-HIQ 1129
1063 IDRALGIPDLVSGQALEPQDPSYOPQSPVMIIDQKRPYHGYAGQAMASAGGFNNMQ 1122
1130 GQSPFNSMNMOMN-QQGNFPLQGMHPRANIMPRNTKPKOLMLOQLQOQOPLNQR 1188
1123 GQHPPTVWGMNQOQGMHPRANILIRNNNIFROLMLQOQOQOPLNQR 1182
1189 QALEKMENTAGGAVMRPMQ---PQGFLLAQVAVORSRELSHFRQORVAMMMQ 1245
1183 QALEKMENTAGGAVMRPMQ---PQGFLLAQVAVORSRELSHFRQORVAMMMQ 1240
1246 QO 1305
1241 QO 1284
1306 YGNGQDPDAFGVSSPPNANMSSRMGSONPMQHPQALSIQSSMKGSPGSLARNS 1365
1285 YGNGQDPDAFGVSSPPNANMSSRMGSONPMQHPQALSIQSSMKGSPGSLARNS 1341
1366 SFQSOQFAGHGNPARYSMVHNSGSGHMGOMNNPMPMSGMPGPOKXC 1415
1342 SFQSOQFAGHGNPARYSMVHNSGSGHMGOMNNPMPMSGMPGPOKXC 1391

RESULT 8
US-09-125-635-8
Sequence 8, Application US/09125635
GENERAL INFORMATION:
APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
FILE REFERENCE: 49944
CURRENT APPLICATION NUMBER: US/09/125,635
CURRENT FILING DATE: 1998-08-21
PRIOR APPLICATION NUMBER: 60/049,728
PRIOR FILING DATE: 1997-06-17
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 8
LENGTH: 951
TYPE: PRT
ORGANISM: Homo sapiens
US-09-125-635-8

Query Match 66.9%; Score 4938; DB 15; Length 951;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 950; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

300 CIOFRLNDGOSMSOKRHOEAYLNGHAETPYRSLADGTIVTAQRTSKLFRNPVTND 359
1 CIOFRLNDGOSMSOKRHOEAYLNGHAETPYRSLADGTIVTAQRTSKLFRNPVTND 60
360 RHGFVSTHFLQEQNGYRPNPVPVGGIRPMAAGCNSSVSGMSMSPNOGLQMPSSRAYGL 419
61 RHGFVSTHFLQEQNGYRPNPVPVGGIRPMAAGCNSSVSGMSMSPNOGLQMPSSRAYGL 120

Qy	420	ADPSTGOMSGARXGSSGNINSLTFCGGMGOSPSTYONNNVGLNMSSPSPHSGPLAPMOQ	479
Dp	121	ADPSTGOMSGARXGSSGNINSLTFCGGMGOSPSTYONNNVGLNMSSPSPHSGPLAPMOQ	180
Qy	480	IMISPRNNGSKRIASHOFSPVAGVHSPMASSGNTGNHSSFSSSLSALOAISEGVGTSLLS	539
Dp	181	IMISPRNNGSKRIASHOFSPVAGVHSPMASSGNTGNHSSFSSSLSALOAISEGVGTSLLS	240
Qy	540	TLSSPGKPLDNPMMNNTTQPSKVSYNODSKSPFLGYCQONPEVSSMCOSSNRDHLSDKESK	599
Dp	241	TLSSPGKPLDNPMMNNTTQPSKVSYNODSKSPFLGYCQONPEVSSMCOSSNRDHLSDKESK	300
Qy	600	ESSVEGAONONGPRLSENGHKLLKLLTLCSSPDRRHSSLTSPDSSCKESSVATSPSGV	659
Dp	301	ESSVEGAONONGPRLSENGHKLLKLLTLCSSPDRRHSSLTSPDSSCKESSVATSPSGV	360
Qy	660	SSSTSGGVSTSNMHSGLDCKHRILKLLQNGNSPAEVAKITAQAATGKPTSSITSCGDG	719
Dp	361	SSSTSGGVSTSNMHSGLDCKHRILKLLQNGNSPAEVAKITAQAATGKPTSSITSCGDG	420
Qy	720	NVVKOEQLSPKKKENNALLRYLLDRDDPSDALSKELQPOVEGVDNKMSQCTSSITPSSSQ	779
Dp	421	NVVKOEQLSPKKKENNALLRYLLDRDDPSDALSKELQPOVEGVDNKMSQCTSSITPSSSQ	480
Qy	780	EKDPKIKETSEBESGDLNDALITGLDTSDFNNSSISSNGSHLTGKQOVPQGTNSLGL	839
Dp	481	EKDPKIKETSEBESGDLNDALITGLDTSDFNNSSISSNGSHLTGKQOVPQGTNSLGL	540
Qy	840	KSSQSVOSIRRPYRRAVSLSPSVSGSSPPYKNISAPPLMKOPMLGANGRMMDSQENYG	899
Dp	541	KSSQSVOSIRRPYRRAVSLSPSVSGSSPPYKNISAPPLMKOPMLGANGRMMDSQENYG	600
Qy	900	SSMGGPNRNTVTQTTPSSGDMGLFNSTAGRMPEMNSNSMKGPRGQDVTSLPRPALGGSIP	959
Dp	601	SSMGGPNRNTVTQTTPSSGDMGLFNSTAGRMPEMNSNSMKGPRGQDVTSLPRPALGGSIP	660
Qy	960	TLPLRSNSTIPCARPYLQOQOQOMLIMRGEITPMGGMANPYCOGAASNOJGSPMGMLSMEO	1011
Dp	661	TLPLRSNSTIPCARPYLQOQOQOMLIMRGEITPMGGMANPYCOGAASNOJGSPMGMLSMEO	720
Qy	1020	VSHGTONRPLLRNSLIDLTVGPPSNLEGOSEDERALLDOLHTLLNSTWDTAGLEEDRALGIP	1071
Dp	721	VSHGTONRPLLRNSLIDLTVGPPSNLEGOSEDERALLDOLHTLLNSTWDTAGLEEDRALGIP	780
Qy	1080	ELVNOGQALEFKODAFQGOEAAVMMDOKAGLYGQTYPAQCGPRMGOGFNLGOGSPSPNSMM	1133
Dp	781	ELVNOGQALEFKODAFQGOEAAVMMDOKAGLYGQTYPAQCGPRMGOGFNLGOGSPSPNSMM	840
Qy	1140	NOMNOGQFRLQGNHPRANTIRPTNTPKOJRMLOLQORLQOGQFLNOSRALELKMNPT	119
Dp	841	NOMNOGQFRLQGNHPRANTIRPTNTPKOJRMLOLQORLQOGQFLNOSRALELKMNPT	900
Qy	1200	AGGAAYVMPRMQPOQGFLLAQMVAQORSRELLSHHFROQRYAMMMQOQOQOQO 1250	
Dp	901	AGGAAYVMPRMQPOQGFLLAQMVAQORSRELLSHHFROQRYAMMMQOQOQOQO 951	
RESULT 9			
US-08-891-640-2			
Sequence 2, Application US/08891640			
GENERAL INFORMATION:			
APPLICANT: Chambon, Pierre			
APPLICANT: Gronemeyer, Hinrich			
APPLICANT: Voegel, Johannes			
APPLICANT: Lutz, Yves			
TITLE OF INVENTION: Transcriptional Intermediary Factor-2			
NUMBER OF SEQUENCES: 14			
CORRESPONDENCE ADDRESSES:			
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.			
STREET: 1100 New York Avenue, NW, Suite 600			
CITY: Washington			
STATE: DC			
COUNTRY: USA			

```

1 ZIP: 20005-3934
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #1.0, Version #1.30
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/08/891,640
16
17 FILING DATE: Herewith
18
19 CLASSIFICATION: 435
20
21 PRIOR APPLICATION DATA:
22
23 APPLICATION NUMBER: US 60/021,247
24
25 FILING DATE: 12-JUL-1996
26
27 ATTORNEY/AGENT INFORMATION:
28
29 NAME: Steffe, Eric K.
30
31 REGISTRATION NUMBER: 36,688
32
33 REFERENCE/DOCKET NUMBER: 1383_0130001/EKS
34
35 TELECOMMUNICATION INFORMATION:
36
37 TELEPHONE: 202-371-2600
38
39 TELEFAX: 202-371-2540
40
41 INFORMATION FOR SEQ ID NO: 2:
42
43 SEQUENCE CHARACTERISTICS:
44
45 LENGTH: 1464 amino acids
46
47 TYPE: amino acid
48
49 TOPOLOGY: linear
50
51 MOLECULE TYPE: protein
52
53 US-08-891-640-2

```

[illegible]

[illegible][illegible]

QY 1205 VBRPMPOGFLNAQVARSRELLSHPRQOAVAMMQ0000000000000000 1264
DB 867 TLRPGV-PTQAPINAOMLQROREILNOHLR-----0000000000000000 911
QY 1265 QOQTOAFSPPPNTASPMGLAGPTMPOAPQOEPYQPNYMGQOPPAFGVSSPPN 1324
DB 912 MMRGGLNTPMSVAVSGMATSMNPRIPQANNOQFPFPNPGISQOPDGFCAITPOS 971
QY 1325 AMMSRMGSGQNMQHPQOASTYOS-SEKKGPMSCNLAIRNSFSQO---QFAHQGNPAY 1380
DB 972 PLMSPRMATHOSQMOOSQANPAYQAPSDINGMAGNMGNSMFSQOSPFPFGQOANTSM 1031
QY 1381 YS-----WYHNGSSGHMGOMNNMPPMPCGM-DMGPDQ 1412
DB 1032 YSNMNNINYSMATNTGMSMNMOTGQISMTSVTSVPSGLSGMGPBQ 1079

RESULT 12
US-09-125-635-2
Sequence 2, Application US/09125635
GENERAL INFORMATION:
APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
FILE REFERENCE: 49944
CURRENT APPLICATION NUMBER: US/09/125,635
PRIOR FILING DATE: 1998-08-21
PRIOR APPLICATION NUMBER: 60/049,728
PRIOR FILING DATE: 1997-06-17
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 186
TYPE: PRT
ORGANISM: Homo sapiens
US-09-125-635-2

Query Match 13.0%; Score 957; DB 15; Length 186;
Best Local Similarity 98.4%; Pred. No. 5,4e-56;
Matches 183; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 117 LQALDGLFVVRANIVFSENVTOYLOQKOEDLVNTSVYILHEEDRKDLKLNPKS 176
DB 1 LQALDGLFVVRANIVFSENVTOYLOQKOEDLVNTSVYILHEEDRKDLKLNPKS 60
QY 177 TVNGVSWTNEPQOKSHTECMRLMKTDPHDLIEDINASPMPORRYETMQCFALSOPRAM 236
DB 61 TVNGVSWTNEPQOKSHTECMRLMKTDPHDLIEDINASPMPORRYETMQCFALSOPRAM 120
QY 237 EEEEDLOSCMIVARRITGERTFSPNPEFTRHDLSCGVVNIIDNLSMRPGEFI 236
DB 121 EEEEDLOSCMIVARRITGERTFSPNPEFTRHDLSCGVVNIIDNLSMRPGEFI 180
QY 297 IRRCIQ 302
DB 181 IRRCIQ 186

RESULT 13
US-08-891-640-3
Sequence 3, Application US/08891640
GENERAL INFORMATION:
APPLICANT: Chambon, Pierre
APPLICANT: Gronemeyer, Hinfich
APPLICANT: Voegel, Johannes
APPLICANT: Lutz, Yves
TITLE OF INVENTION: Transcriptional Intermediary Factor-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW, Suite 600
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,640
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/021,247
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Steife, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1383.0130001/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1036 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-891-640-3

Query Match 11.8%; Score 873; DB 12; Length 1036;
Best Local Similarity 27.7%; Pred. No. 4.1e-49;
Matches 321; Conservative 173; Mismatches 342; Indels 322; Gaps 54;

QY 378 PNPNP-VGGIRPPMAGCNSVCGMSPNQGLQMPSSRAYGLADSTGQMSGARYGGS 436
DB 4 PRNPNVSNBSIPA-----HGVARSSTLPPSNMNVSTR-----INQOOSDILSSHSNS 54
QY 437 SNI-ASTLPQGMQ-----SPSSYONNNYGLNMSSPPHSGPGLAPNQONIMISPRN 486
DB 55 SNQSGFSGCPSGQIYANVALNKGAOSSQSKPSLNLNPPMGCTGISLAQ---FMSPR 111
QY 487 RGS-----PKASHQSP-VACVHSPM---ASGATGHSFSSSLALQAISEVGTG 536
DB 112 QVTSGLATRPMPNNSPPISTLSPVGMTSSACNNNNNSYNIPVTSLOGNNEGPNN 171
QY 537 LESTLSP---GPKLNSPMNITOPSKYNSODS---SPL-----GFCQDNVESS 583
DB 172 VGFSSASSPVLRQMSONSBSRLNIQPAKASKKNEKIASTLNMQSDNSSSGKPLDSG 231
QY 584 MCSNSRDLHLSDEKSESSVEGAENQGPLSKGHRKLLQLTCCSSDDRCHSLTNSPLD 643
DB 232 LHNHN--DRLSDGDSKYSQS-----HKLVQLLTFTAQO---LRHADID 271
QY 644 SSCRESSVYTPSPGVSSSTSGVSVSTSNMHSGLQEKHRLKLLQNGSPAFAKITA 703
DB 272 TSCKD-VLSCGTGNSASANSAGSCPS--HSS-LTARHKLRLHLOEG--SPSDITTLV 327
QY 704 QATGKPTSTSCGDDNV-----VKOEQLSPKKE--NNALLRYLDRD-----PSDA 750
DB 328 EPPKDSASTSVYTOQOVGNSSIKLELDASKKESKDHQLLRLDKDEKDLRSTPNNS 387
QY 751 LSKELQPOVEGVNKMKSQCTSTIPSSQEKDKIRKTTSESGSGDLDNLAILGDLTSS 810
DB 388 LD-DVKVKEKE-QMDPCNTINFTPMKATPE-EIKLEAQSGTADLDQDQL----- 438
QY 811 DFTNNSTSSNGSHLGTRQOYFGQTSNLGLKSSQSVOSIRPPYRAVSLDPSVSGSPPY 870
DB 439 ----- 438
QY 871 KNISAFPMLEKQPMGLGNPRMMDQENYSGSSMGPBRNVTVTOTPSGDMGLPNSKAGRN 930

```

Db 439 -----P1LEKAAQL-----PGICETDR-----MGCAVTSVITKEILITKEILPASTL-----480
Qy 931 EPMSNSMKRPGGDYNTSLPRALGSIPLPLRKSINPGARVYLQOQOOLQMRPEIT 990
Db 481 -----OSATARP-----TS-----RLNRLP-----EDE 498
Qy 991 MGMGANPRYGOAAASNOIGSMPPDCLSMEOVSHGTQNRPLLRNSLDLDVGRPSNLEQSD 1050
Db 499 LEAIDNOEGOGCTGDDI-PWNTNN--TYTAINOSKSEDDCISOLDDELCPPTTVBERND 555
Qy 1051 RALIDHTLLTSNTDAGLEIBRAGIPLVLNMOGALFEKQAFQGOEAA--VMDQKA 1100
Db 556 KALLEQVLSFGKDEIHELALDRALGIDKLV-QGGGLDVLSRFPPQCATPLLINEEP 614
Qy 1109 GLVQYTPACGP-----PMGGFHLQGGSPFSFNSMMQNN-QQGNF-PLQGHPRANTM 1160
Db 615 NLYSQPYSPPTANLSPSPG--WVRQKPSLGTMPVQVTPPRGAFSPGCMQPROTLN 671
Qy 1161 RPRNTFKQLRMQLOQLQLOG-QQFLNOSRQALELKMBNPLAGAAVYRPMQO--QGF 1210
Db 672 RPPA-APNOUHLQLOQRLLQGOOOLIHONRAI-LNOFAATAPRGVIMKRSGMDOQITP 729
Qy 1217 LNAQMVAORSRELLSHHFRRQRVAMMMQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 1270
Db 730 LNAQMLQORQRELYS-----QQHRQOLIQORRAMLMQKQSGNNLP 773
Qy 1277 VTASPSMDGLAGFTMPQAPPOGFYQPNYGMQOQDPDA---FGHVSPPNMASSR- 1330
Db 774 -SGLEPQGT--NRLRFGAPRQGFYPRYNTGTMPGTTPASTSPFQLANPEASLANRS 829
Qy 1331 -----MGPSQNPMMQH-----P 1340
Db 830 MYSKGMTGNIGOGFTGINTQMOQNVFYQYAGMVFQGEANFAPSLSPSSWVMPDIPP 889
Qy 1343 QA-----ASTYOSSEMKGWPISGLNARSSFS--QQAHGQNPAYS--WVHNQSS 1390
Db 890 QSSLLQCTPPASGYSQDMMKAMQOGALGNVNFSAVQNPQAPQGVYNNMSTIVSM 949
Qy 1391 GHMGOMNNMPKMSGMP 1408
Db 950 GNTVQNMNPM-MAQMQM 966

RESULT 14
US-08-701-154A-5
: Sequence 5, Application US/08701154A
: GENERAL INFORMATION:
: APPLICANT: O'Malley, Bert W.
: APPLICANT: Tsai, Ming-Jer
: APPLICANT: Tsai, Sophia Y.
: APPLICANT: Onate, Sergio A.
: TITLE OF INVENTION: STEROID RECEPTOR COACTIVATOR
: TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FASTSEQ for Windows 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/701,154A
: FILING DATE: August 21, 1996
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:

```

```

      APPLICATION NUMBER: 60/003,784
      FILING DATE: September 15, 1995
      ATTORNEY/AGENT INFORMATION:
        NAME: Warburg, Richard J.
        REGISTRATION NUMBER: 32,327
        REFERENCE/DOCKET NUMBER: 220/243
      TELECOMMUNICATION INFORMATION:
        TELEPHONE: (213) 489-1600
        TELEFAX: (213) 955-0440
        TELEX: 67-3510
      INFORMATION FOR SEQ ID NO: 5:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 1061 amino acids
        TYPE: amino acid
        STRANDEDNESS: single
        TOPOLOGY: linear
      US-08-701-154A-5

Query Match          11.8%; Score 872; DB 11; Length 1061;
Best Local Similarity 27.1%; Pred. No. 4,9e+49;
Matches 316; Conservative 174; Mismatches 334; Indels 340; Gaps 52.

QY   378 PNPNP-VGGGIRPPMAGCSSVYGGMSSMSNOGLMPSSAYGLADPSTTGQSGARYGS 436
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    4 PPNVSVPNSISPA-----HGVARSTLTPPSNSMVSTR----INQQSDSHSSHSHS 54

QY   437 SNI-ASLRPGPMO-----SPSYONNNYGLMSSPPHSGPEGLAPNQCINIMSPN 486
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    55 SNSQSGFGCPSQIYANVALNKGAQSQSSKPSLNLNPPMEGTGISLAQ--EFSPR 111

QY   487 RGS-----PKLASHQFP-VAGVHSPM--ASSGNTGHSEFSSSLALQAISEGVTS 536
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    112 QVTSGLATRPRRPNNSFPNISTLSPPGMTSSACNNNRYSNLPVTSLOQMNEPNS 171

QY   537 LLISSSP---GPKLDSNGNMNITOPSKYSNODSK---SPL-----GFYCDQNVESS 583
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    172 VGFSASSPYLAROMSSONSFPSRLNIOPAKAESKDKNELASTLEMIQSDNSSDGKPLD 231

QY   584 MGOSNRHLISKESKESEVEGAENQRPLEKGHKILQLLTCSDDRGHSLLNPSLD 643
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    232 LIHNH-DLISGDCKYSQTS-----HKLVQLTTTBEOO---LRHADTD 271

QY   644 SSCKESSVSVTSPGYSSSTSGVSSSTSMMHSLDLQEKRIHLKLIQNGNSPAEVAKITA 703
      : | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    272 TSCKD-VLSCGTGSNSASANSNGSGSPSS-HSS-LTARKKLIARLQDE-SPSDITTL 327

QY   704 QATKDTSSITTCGDCNV-----VKOEQLSPKKK--ANNALLRYLIDRD-----PSDA 750
      : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    328 EBDKDKDSASTSVYGOVGNSSIKILELDAASKKESKHQLLYRLLDKDEKPLRSTPNLS 387

QY   751 LKEKLOPEGVGNMKSOCTSTPISSSQEKPKITETSEEGSGDLMDLIDLGLTSS 810
      : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db    388 LD-DKAVKYEKE-QNDPCNTNPTPTKTATPE-EIKLEQSOFTADLDQFDOLL----- 438

QY   811 DEYNSSISSNGSHLGTKQCVFOGTNSLIGKSSOSVOSIRPNRAVSLDSPYSGSSPV 870
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    439 ----- 438

QY   871 KNISAFPMILPKCPMLIGNFRMMSOENYSSMGCPNRNVTVTOTSSGDMGLPSNKAGM 930
      : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    439 -----PTLEKAAOLPG-----LCETDEM 456

QY   931 EEMNSMSMRPGGDVNTSLPRALGSIITPLRLSNSIFG-----ARRVLQOOQOOLQM 984
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    457 D-----GAVTSVIKSELIPASLASQATAARTSRLN----- 486

QY   985 REGETPMGANGPYCOAASNOLGSPWDMLEMEQVSHQTQNRPLRLNSLDLVGPPSNL 1044
      : : : : | | | | | | | | | | | | | | | | | | | | | | | |
Db    487 RLPELELELDNQFQPGTGDGI-PMTNN--IVTAINQKSSEDQCISSLQDELICPPTV 543

QY   1045 EGQSDERRALLDQLHTLLSTDATGLEEDRALGIPELYVNOGALBPKODAFQGEAA--V 1102
      : : : : : : | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 544 EGRNDEKALLEJLSEFLSGKDETELAEIDRLGIDKLV--QGGGLDVLSEFPPOQATPPL 602
QY 1103 MMDQKAGLVGQYTPACG-----PMGGFHLQGSPPFMSMNQAN--QOGNF--PIQGNH 1154
Db 603 IMERNLTSQPYSSPPFANLPSPPQG--MYRQKPSGLTMVQYTPRGAFSPGGMQ 659
QY 1155 PRANIRPPTNTPKQRLMOLOQRLQ--QOFLNQSROALELKMENPTAGCAAVRPMMPQ 1213
Db 660 PROTLMRPPA-APNQLRLQQLQQLQOQQLIHONRAI--LNQFAATAPVGINNRSGMQQ 717
QY 1214 ---QGFNAQMAVQORSRELLSHHFROQRYAMMMQOQOQOQOQOQOQOQOQOQO 1270
Db 718 ITQPPLNQMQLAQORQRELTS-----QQHRQRLIQOQRAMLMRQGSFG 761
QY 1271 FSPPPNVTASPSMDGLLAGPTMQAPQPPQYPTNGMQQPPA----FGRVSSPPNM 1326
Db 762 NNLPSP--SGLPVQGT--NRRLPQAPQOFPYPPNYGTNPCTPASTSPSOLAANPEAS 817
QY 1327 MSSR-----MGPSQNPMMQH----- 1341
Db 818 LANRNSMVRGKMTGNIGGQGTGCTINPQMOQNVQYPGAGNVPGCANFAPSLSPGSSMP 877
QY 1342 ----PQA-----ASTYQSEMKGWPSSGNIARNSSFSQ--QOFAHQGNPAVYS--MY 1384
Db 878 MP1PPQSSILQOTPPASGVQSPDKAMQGAIGNNNVFSQAVQNPPTPAQPGVYNNMSI 937
QY 1385 HMGSSGCHGCMNMNMMPMSGMP 1408
Db 938 TVSMAGNTNVQNMNM-MAQOMQ 960

RESULT 15

US-60-236-359-19468
; Sequence 19468, Application US/60236359
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: MDMORE-4P
; CURRENT APPLICATION NUMBER: US/60/236,359
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 60/207,456
; NUMBER OF SEQ ID NOS: 21/09
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 19468
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034418.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: AM503706.1, EVALUOE 9.00e-51
; OTHER INFORMATION: SWISSPROT HIT: Q58434, EVALUOE 5.30e-01
US-60-236-359-19468

Query Match 10.1%; Score 742; DB 23; Length 149;
Best Local Similarity 99.3%; Pred. No. 8.6e-42;
Matches 148; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 644 SSKESVSVTSPSGVSSSTSGGVSSSTSNMHSLLQEKHRIHLKLLONGNSPAEVAKITA 703
Db 1 SSKESVSVTSPSGVSSSTSGGVSSSTSNMHSLLQEKHRIHLKLLONGNSPAEVAKITA 60
QY 704 QATGKDTSSITSCGDNVYKQEQLSPRKKENNALRYLLDRDPSDALSKELQPOVEGYD 763
Db 61 EATGKDTSSITSCGDNVYKQEQLSPRKKENNALRYLLDRDPSDALSKELQPOVEGYD 120
QY 764 NKMSQCTSSITPSSSQEKDPKITEETSEE 792
Db 121 NKMSQCTSSITPSSSQEKDPKITEETSEE 149

Search completed: April 19, 2001, 16:54:00
Job time: 487 sec

[illegible]

Query Match	2.6%	Score 189.5	DB 5	Length 849
Best Local Similarity	17.8%	Pred. NO. 2.3e-05		
Matches 169	Conservative 147	Mismatches 206	Indels 337	Gaps 39
QY	577	ONPVESSMCQ---	SNSRDLSDKESKESSEVEAENRGPLESKHKLLQLLTCSDDRG	633
		: : :	: : :	: : :
Db	6	EADTDESTANDLYQOQNEHETFODKKXKHQVEKOSSEETRG--	1HVTDSNDDIG	62
QY	634	-----HSSLTNSPLDSSCKESSVSVTPSPGCVSSSTGCGVSTSNMGSLLQEKHRI--		684
		: : : : : :	: : : : : :	: : : : : :

Db	63	VTXSODTEEVLEIENSFPVDOLKEQOEESTFKMS-----LSERDLVDE	102
Qy	685	LHKLLQN-----GNSPAEYAKITTAQATKDTSSITSCGDNVNVKEOLSPKKRENN--	735
Db	103	IDELPTNSTKTYENNGQSEITNKRAYE-----SVETPOELTPMDKROKLD	147
Qy	736	-----ALLRYLLDRDDPSDAL--SKELOP-----OVEDYKMSQ	768
Db	148	ANTEITSVTELESVNNHNEQSOPIEPQOEOPSTTETTYTSISVPVSTTNVEKASSINE	207
Qy	769	CYSSSTIPSSQEKDKIKTERSEEGSD-----LONLAILGDLTSSDFYNNSSISNG	821
Db	208	QEDLEMIAKYQOATINLEITERAMEGHGHDGCHFTOEN-----GQPSGSLISSIYPSDS	262
Qy	822	SHLGTKQ-----OVFOGTSNLGKSSQVOSIRIPRYNRAVSLDSPY-----S	863
Db	263	ELLNTNQVAAVATYLSLSQLEQHTSASAMLSATLSAL-----PLSTIAVYLPRIQLL	316
Qy	864	VQSSPVPVNIASFPLRPQPMGLGNPRAMDQENYVSGSMGKPRNNYV----VYQTPSSGD	919
Db	317	INTLPTLONLNTV--OLLEFVATSPYQKIIDLASPDPTSAQATYKRDITSLEFETIKRUYSED	374
Qy	920	WGLPNSKAGREPMNSNSMGRP-----GDYNTSLPRPLGSLIPTLPLRSNISPAGRPVL	975
Db	375	-----DPLFTEVHIARGMKMBEEDRSTLFKQKQSIESTLKKVLAFLFATLIG	423
Qy	976	QOQOQOMLOMRGELRPMGKANPYGOAALASNOUGSPMDGMLSMEDVSHGTGNRPPLRNSLD	1035
Db	424	TMEIGFFYLINESFLDFVCPSSNNLDPNSALMSIGYONGLOSTDS--PYGARVYGKLK----	478
Qy	1036	DLVGPPSNL-----EGQSDERALL-----DOLH-----	1058
Db	479	---PQATLYLDIKQATISAIEAGERKEKEILEDILPDDLHYLMSRRNAKLLSPETD	534
Qy	1059	-----TLNSTDATGLEB-----	1071
Db	535	FVMRCKORKESSLNTTETETPLSEQYDWFTEPLRLEFDVYSKNIAVLIWGMKGMTKMKNRRED	594
Qy	1072	-----IDRALGRIPELYNO-----GOA-----LEP--KQDAPQOGAAVMMQOKAGLYG	1112
Db	595	TPHPOELLDNTTGSTQPMQNLSSSSGQASSTPYVDPRKMLYSMRKANIANVKK-----	648
Qy	1113	QTYPAO-----GRPMQGFHLOGQSPFSNSMMNOMNOGNEPLOGMHPRANINMPT	1164
Db	649	---PQORAWMSREEEKALRIHALETGK--PHMATILTEFGGCKI-----SEALKNRT	695
Qy	1165	NTRKQLRMNQLOLRLOGQFGLNMQALEL-----KMENPTAGAAVMMP	1208
Db	696	-----OVQJLKDXARNNKKFEF---HSGLEIPSYLGRVGTGVDDGKRKKDKNTYKTTAAAPV	748
Qy	1209	MMQPOGGLNMQVAVRRELISSHFRQVAVM-----MMQOQOQOQOQOQOQOQOQOQOQOQ	1264
Db	749	NMSEQ-----LQOQOQROQEKQKQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ	803
Qy	1265	QOQOQATSESPRPVNTASPSMDGLLAGPTWPCARPQOQFFYQPNYMGQOQPD	1313
Db	804	AEQEO-----PQOPEEQOQOQEKQEPQO---QOPD	829

```

RESULT      3
US-60-248-505-1205
; Sequence 1205, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1205

```

LENGTH: 1355
 TYPE: PRT
 ORGANISM: Human
 US-60-248-505-1205

Query Match 2.3%; Score 166.5; DB 6; Length 1355;
 Best local similarity 17.9%; Pred. No. 0.0012;
 Matches 265; Conservative 111; Mismatches 510; Indels 535; Gaps 64;

QY 10 PLASDRKRLPCDTPGQ--LTCGSEKRRREDE-SKYEIELLAELISANLSDIDNENKPK 66
 DB 211 PMAG-SKTSSTPERKODKEISOEGDEPARQSVSKIRQFGRQ-BGNLA----- 259
 QY 67 DKCAILKETVROIRKIQEOKTISNDDDVAKADVSSVTCGVGIDKDSIGPLLLAALDGLF 126
 DB 260 TQSSPPEAVQRCPEDETKSTDSKDVCRMPDQEPGKADQTPATKMLGEPED--- 315
 QY 127 VVNRANIVFSENVTOYLQKQEDL-VNTSVYNIILHEEDKDFLKLMPKSTVNGVSWTN 185
 DB 316 -----YGRSTQERKCEIKDLFVQYGSRRGSETSDMRDERKRRRQPEAHGTAGQK 366
 QY 186 EPQRQSHTEFCRLMKTPHDILEDINASPQRQRYETMOCFALSDPRAMEGEDLQSC 245
 DB 367 ERDRKWL-----VLE-----TOTDCKYQELQGLSKSKDAEKSETOYLLS- 407
 QY 246 MICVARITIGTERFP-----SNPESFITRDLGSKVYNIDTNSLRSSMRGFEIIR 298
 DB 408 -----EGGDDTHPELEGTAVSGEAEHTKEGTAEAFVN-----SKNAPAAE- 448
 QY 299 RCIOREFSLNDGQSWQKRYHQAETPYRFSIADGTIVTAQTSKLFRNPVTN 358
 DB 449 -----KTLGARETQDL-----APLEKQSVGEINRYT-----KTHDOYV- 482
 QY 359 DRGFEVSTHLOREQNCYRPNPVPVGGIRPPMAGCNSSVGVGMSMSNOGLQMPSSRAYG 418
 DB 483 -----EEDDQY-----QG-EDPESPFTQSDGSESETPNS----- 510
 QY 419 LADPSTTGQMSGARVYGGSSNIASTLTPGPMQSPSSYONNNG-----LNMSSPHGS 470
 DB 511 -----LASEEGSSSETGELP---VOGDSQSGDGHGSEVQGGHNNNPDTQROGT 557
 QY 471 PGLAPNOQINIMISPRNG-----SPKIASHQ-----ESPVAGVH---SPMASSGNT 513
 DB 558 PG-EKNALALAVYPVAVGEVQVLTDEQEPREHKNQGPRTKGRPAVAERPHGPEAQSTA 616
 QY 514 GNSHSSSSLSALQAISEGVSTLSTLSSPGPKLNSPMMNITTOPSKVSNQDSKSPILGF 573
 DB 617 GDNKRKSLTEITGALDEDF-TDOLSLMQLPG-KGDSRNELKVOGFP----- 661
 QY 574 YCQONPVESSMCOGNSNDHLSDKESKSSVEGAEENQGPLESKGHKLLQLLTCSDDRG 633
 DB 662 -----SKEKGRATEAO----- 673
 QY 634 HSSLTNSPLDSCKSSSVTSPGVSSTSGVSSSTNNHGSLLQEKHRLHLKLLONGN 693
 DB 674 -----HTLESLEDBDNASLX-----IOLKTEPVTSE-----EED 705
 QY 694 SPPEVAKITQAQTKDTSITSCGDGNVVOEOLSPKPKKNNALLRLYLLDRDPSDALSK 753
 DB 706 SPOELA-----GEGG-----DQKSPAKKEHNSV-----PMSLEK 736
 QY 754 ELQPOVGVNKKMSQCSSTIIPSSQEKDKIKITESEBSGDLNLDLILGLTSDFY 813
 DB 737 QMRDQPCSVGVAVYSSPLVYLOLKIQQTNVQEEHQKQV-----IAQASGELC 791
 QY 814 NNSISSNGSLGTQOQFOGTNSLGLKSSQSVOSIRPPYRAVSLD-SPVYSGSSPPV- 870
 DB 792 SVSLTSLISDC-----SVFFNYSAQO-----PYTRGLPLDPSRPAQOETPAPQA 836
 QY 871 ---KNIS-----AFPV-----LDPKOP 883
 DB 837 LEDKOVSSVHNSLKMNSLSVFCSWFRRRSWPCRHOPACLVREAFAAGRAHAPAPAVPR 896

QY 884 MLCGNPRAMDSENYGSSM-----GGPNRVTV---TQPS-----SGDWGLPNSKAGME 931
 DB 897 IYGRFFLEFNQRHLGFSFVYKMDGAPMLCLIPRNTGTQORVLPVYVSPSKKRYLS 936
 QY 932 PMNSNSMG-----RPGDDYNTSLP---RPAIGSIPPLPLRSNSIPGAPVLAQOQOQ 980
 DB 957 ACNSMMGHLSPVRIPLRCKFNQLPRLSDQVYIPALPLKMEVRAEERKATEVKQDQVET 1016
 QY 981 MLOMRGEIIPMGKANDPYGAALASNOL-----GSPDGMUSMEOVSHGTONRPLRN 1032
 DB 1017 Q-----GOENKRGKPCSCNGEAASTSLFTQGNLTSSWYN----- 1051
 QY 1033 SLDDLVGPNSLRGQSDERLLD-----QHLTL-----LSNTDQGLEET--- 1072
 DB 1052 -----PRLGCVHKLSTLEKNQTDKAQVHAVSFYSKDHEVASHSPAGGILPFGPK 1103
 QY 1073 -----DRALGIPELVNOGOALEPRKDAFQO-----Q 1098
 DB 1104 DPLPTVLPAPVPGCSLMPKRAALKVLOKDLHPSSPGLMVGEDMQPKDPALASSRSSPS 1163
 QY 1099 EAAVYMDQKAGLYGQYPAQGPMPQGFHLQGSFSPNSMMNNOQG---NFPLOGMH 1154
 DB 1164 RAASHSSHKRKL-----SEPLQ-----LQPTFP---LQKMDRDEGPPPAKFPCC--LS 1207
 QY 1155 PRA---NIMRPTNTPKOLRMQLOQLQOQFL-----NOSRALELKMENPT 1199
 DB 1208 PEALLTERLKKRKNKTLIMVLAESKSKRESGETVYVKEVLDIKKWRBDPCKDKOKDOK 1267
 QY 1200 AGCAAVMRPMQPOQGFILANQVNAORSRELSHHFROQRAVMMQOQOQOQOQOQOQ 1259
 DB 1268 AGNOTKKKROOTDQ-----SRESODIMRHVVRQ-----IPTPDDSHGVKAHNOE 1314
 QY 1260 QOQOQOQOQAFSPP-----NVTAS---PSMDGILLAGPT 1291
 DB 1315 SQRIRRHQVRLDSPEONIGRHQVTGOAQIPNSNGAKGET 1355

RESULT 4
 US-09-649-996-14
 Sequence 14, Application US/09649996
 GENERAL INFORMATION:
 APPLICANT: Jomo, Keith E.
 TITLE OF INVENTION: KINASE GENES AND USES
 NUMBER OF SEQUENCES: 53
 CORRESPONDENCE ADDRESS:
 ADDRESS: Lyon & Lyon
 STREET: 633 West Fifth Street
 Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Fastseq for Windows 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/649, 996
 FILING DATE: 29-Aug-2000
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/976,255
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 229/182
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600

```

1 TELEFAX: (213) 955-0440
2
3 TELEX: 67-3510
4
5 INFORMATION FOR SEQ ID NO: 14:
6
7     SEQUENCE CHARACTERISTICS:
8
9         LENGTH: 1503 amino acids
10
11         TYPE: amino acid
12
13     STRANDEDNESS: single
14
15     TOPOLOGY: linear
16
17     MOLECULE TYPE: Protein
18
19     SEQUENCE DESCRIPTION: SEQ ID NO: 14:
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
10
```

Query Match	2.28	Score 165;	DB 5;	Length 1503;
Best Local Similarity	20.98	Pred. No. 0.0017;		
Matches 222;	Conservative 136;	Mismatches 383;	Indels 322;	Gaps 57;

QY	2	SGLENDLPDLSKSRKRLPCD-----TRGGGLCSGCKRRRRRDESNYIEELAL--	51
Db	600	SSTDEDFQOSTDRKDDSLPBDLLVHTSGPESPFFNNINDVDKSDLPSHOKIFDMLNG	659
QY	52	ISANL-----SDIDNFVYKPDKALIKETVROIROIKEGKTISNDDVQKADVSTGQ	105
Db	660	VOADFKPATLSSSLDN-----PRESYITGHFPEKPRKIFDSEPLCD-----	703
QY	106	GVIDKDSLGLLQAL-DGFLFVYNNRENTYFVSENTQYLIQYQEDLVN-----T	155
Db	704	NLMHODNDPLNVELSELNFEFL--QEKNNILKGLSKSEHNIDQTLKNAGFTTEAMLET	761
QY	156	SVYNLHEDRDKDELKLNLPKSTVGVSWTNPEQOKSHTEKRCMLMTPHILIEDIASP	215
Db	762	SCRNSLDTE--LQRAENKP-----GLSLDQNVSTKDDDD--VMLTGDTLSTLOSSP	811
QY	216	EMR---QRETFMOCFALSQPRAMEEGEDLOSCHICVARRITTGERTPPSPESFITRHD	272
Db	812	EVQVPTSPFTEETPRPRVPSLPTQGTOTCLDVIY-----PEDCL-HQD	857
QY	273	LSGRVYND-----TNSL--RSSMRG-PEDII RRCIQRFSLDQSMQKRNHY	320
Db	858	ISPDVTVPVLEILTDAETHSLDNRSQDSPSESEETLR-----LTESDS	901
QY	321	EAYLNGHAETPVYRFSLADGTIVY-----AQTKSLFRNPVYTNDRHGFVSHFLQRE	372
Db	902	-----VLADDLIARVSYGSSLPRLGDELHNKPPSEDDH---SHRLEKN	943
QY	373	QNGRPNPNPVYGGCIRPPMACGNSVGGMSKSPMOGLQMPBSRAYGLADPS--TTGMSG	430
Db	944	LEAVE-TLNQNSKDAKAGLIVALSDDSYSDSLSDLSLAFPPASPSLEFPDLES	1000
QY	431	ARYG-----GSSNIASLTPFGPMQSPSSYQNNNYGLNMPSP-----PHGS	470
Db	1003	VDVIEALLDSGHTPOKLYP-PUKPADSGETEY---NLESPPWTLHPAREGADSEPA	1050
QY	471	-----PGLAPNOQIMISPRNGBSPKIASHOPSPVAGVHSPWASSGNTGNHSFSSSL	523
Db	1058	TTGDGHHGHLPPNRY-IYISDAGG-----HRCETEVTPEFTFAGSGYRDAVFSND	1111
QY	524	SALQALSGV-GTS-----LLSTLSSGPKL-DNSPMNNTTPSKVNSQSKSPKLYTC-	575
Db	1111	SEPERSEVEGTSPSALVLYQEOECPLEPVPLEPS-----AAQDS-----CL	115
QY	576	---DONPVESMCQNSRHDLSIDREKSESVGEAENORGPLESKHK--KILDLTCTSD	630
Db	1154	EARKSQPDEGLSLALNHSDDLRAPEPQATGYRQGVNPTEDBASSPWSVLNALLSSGD	1211
QY	631	D-----RGHS-----SLTNSPLDSSCKESS-----VSYTPSG	658
Db	1214	DFEQDQDRPCTLASTGTNTNELLAYTNSALDKSLSSHSECPKLEKPEIEGKYLGLKLVSG	1270
QY	659	VSSSTGVSYS---TSMNGSLLOEKHRILLKLLQNGNSPRAE--VAKITQANQKOTSSI	713
Db	1274	MLDLSDEQMDADEDENSDDDEDLRAFNHSLSSSESEDETEHPVPIILSNEDGRHRLSL	1330

```

Oy 714 - - - - - SSCGGVNVKQEQLSPPKKENNA - - - - - LPRVLDLRDDPSALSTELOP - - - - - OYEG 761
Db 1334 LKPIAANAPDLPEPDM - - - - - KKEKAVTFPDDVTVYVLFQEPFR - - - - - TKELAPCGGEGACG 1385
Oy 762 VDNK - - - - - MSCOSTSTTPSSSOEKDPKIKETSEEGSG - DLDNLALITDLTYS 810
Db 1386 PDLGAPAPASGSPSLSRICNS - - - - - ESSTDEEGGGEFEMD - - - - - 1421
Oy 811 DFIYNNSTSSNCSHLGTQOQVFOGTNSLGLKSSQVOSIR - - - - - PPINRAVSL - - - - - DSPV 862
Db 1422 DFISSDPFWSK - - - - - TTSLNLSKSLSLTSTSYFSPPPARSTEGSSWPHASPY 1468
Oy 863 SVGSSPPYKNISAPMLPKOPMGLGNPRMMDSG - ENYGSMSG 904
Db 1469 SRFSLISPA-NIASFSL - - - - - THLDSTDEGGSSSEGG 1500

```

RESULT 5
US-09-421-124-44
; Sequence 44, Application US/09421124
; GENERAL INFORMATION:

1
 2 APPLICANT: Futreal, Phillip A
 3 APPLICANT: Wooster, Richard F
 4 APPLICANT: Ashworth, Alan
 5 APPLICANT: Stratton, Michael R
 6
 7 TITLE OF INVENTION: Materials and methods relating to the
 8 identification and sequencing of the BRCA2 cancer
 9 susceptibility gene and uses thereof.
 10
 11 TITLE OF INVENTION:

ADDRESSEE: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh

```

?      COUNTRY:  USA
?      COMPUTER READABLE FORM:
?      MEDIUM TYPE:  Floppy disk
?      COMPUTER:  IBM PC compatible
?      OPERATING SYSTEM:  PC-DOS/MS-DOS
?      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
?

```

```

1      APPLICATION NUMBER: US/09/421,124
2      FILING DATE:
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: US/08/755,587
5      FILING DATE: 25-NOV-1996
6      APPLICATION NUMBER: GB 9523959.6
7      FILING DATE: 23-NOV-1995
8      PRIOR APPLICATION DATA:
9      APPLICATION NUMBER: GB 9525555.0
10     FILING DATE: 14-DEC-1995
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: GB 9617961.9
13     FILING DATE: 28-AUG-1996
14     ATTORNEY/AGENT INFORMATION:
15     NAME: Kenneth D Sibley
16     REGISTRATION NUMBER: 31,665
17     REFERENCE/DOCKET NUMBER: 5405-135
18     INFORMATION FOR SEQ ID NO: 44:
19     SEQUENCE CHARACTERISTICS:
20     LENGTH: 3418 amino acids
21     TYPE: amino acid
22     TOPOLOGY: linear
23     MOLECULE TYPE: protein
24     US-09-421-124-44

```

Query Match	2.28	Score 163	DB 5	Length 3418
Best Local Similarity	17.68	Pred. NO. 0.0067		
Matches 251	Conservative 195	Mismatches 480	Indels 500	Gaps 60

```

0Y      7 NLDPASDSRKRRLPC-----DTPGOGITCSGKRRRQESKYYIEELAEIISAN 55
      ||: ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


Db 1333 NLEFDGSSKNDTVCJHCKETDLLEFTDQHNICKLSGQ-----FMKEGNTQIKED 1393
QY 56 LSDIDNENVAPDKCALIKETVROIRIKEOGKTTISNDDDYOKADVSTGGVIDKDSLGP 115
1384 LSDITFLEV-----AKQGEACHGNTSMKEQDLATKTEQNIKDEFTSPT-----1426
QY 116 LLLQALDELFLVYNREANIVFSEVNTQYLOYKOEDLVNTSVYNIILHEEDRKDELKMLPK 175
Db 1427 -FFOTASGKNISVAKE-----LEFNKIYVFDQKPEELHNSFLNSLSDIRKKNMDILSY 1480
QY 176 STYNGVSWTNEPQROKSHTECNRLMKTPIHDILEDINASEMRORYETMOCFALSQRAM 235
Db 1481 EEDTIVA-----HKILKE-----SVP---1496
QY 236 MESEEDLOSCMICVARRITTEGE--TEPSNPE-----SFTRHDSGKVNIIDNSL 285
Db 1497 -----VGGNOLVTFQOGPERDEKIKEPTLLGFLHASGKVKVIAESL 1539
QY 286 RSSMRGFEEDTIRRCIORFFSLNDGOS-----MSOKRHYOAYLNGHAETPPVRFSL 337
Db 1540 -----DKVKNL-----FDEKEGTSEITSFSHOMAKTLKYREACKD-----LEL 1578
QY 338 ADGTI-VTAQTKSLFRNPYTNDRH-----GFVSTHFLQREON-----GYR 377
Db 1579 ACETITETAPKCKEMONSINNDKNLVIETVPPKLLSDNLKROTENLTKTSKIFLKV 1638
QY 378 PNPYPVQGIIRPPMACGSSVSGMSMSPNOGLQMPSSRAYGLADPST-----425
Db 1639 VHENVEKETAKSP-ATCYTQSPSYTIENSALAFYTCSSKRTSVSOTSLLEAKKMLREGI 1697
QY 426 --GQ---MSGARYG-----SSNIAS-----LTPGQMSPSSYONNNGYLMSSPP 467
Db 1698 FDQPERINADYVGNLYENNSNSTIAENDKNHLSKOTYLSNSMSMSYSVSHDEYV 1757
QY 468 HGSPGLAPNOONIMISPRNGSKRIASHQSPAVGVSPPMASSANTGNHSF-----SS 521
Db 1758 NDRITLISKNLD-----SGIEPVAKNVEDOKNTSFSKVISNVKD 1796
QY 522 SLALQAIISGCVGTSLLSTLSSP-----KLDSPPMNITOPS-----560
Db 1797 ANAYPQVNVNDICVEELVYTSPPCKNNKNAIKLSTISNNFEVGPFPRIASGIRLCSH 1856
QY 561 ----KVSNDOSKPLGFCYDQNPVSSMCO-----SNSRDLSPKES-KESSVEGAENORG 611
Db 1857 ETIKKVIDISFSGVLEFKENNEKSKICQTKMACCYEALDSEDIHNSLDNDE----1912
QY 612 PLSSGKKKLLLOLLTSSDGRHSSSLNSPLDSCKRESSVYSPSSVSSSTSGGV-ST 670
Db 1913 -CSMHSKVFADI--OSEBILQHNQMSGLE-----KVKISIPCDVSLTSDICKCSI 1962
QY 671 SNNHGSLLQEKHRLHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGDGVNVVQEOLESPK 730
Db 1963 GKHLKSY-----SSANTGCIFTASGKS-----VOYSDASLQNAKROVFS 2002
QY 731 KKEN--NALLRYLLDRDPDSDALSKELOVEGVNKMCO-----CTSSITP--SSSQ 779
Db 2003 IEDSTKQVFKSVLEKFSNEHSDQLTRENTAIIRPEHLISQKFSYVNVNSANSRGEFTAS 2062
QY 780 EKRPKIKTIESESGDDLDMLDALGD-----LTSDFYNNSTISSGSH 823
Db 2063 GKQVSTLESLLAKVGLLEFEDLIRTEHSLHSPTSRONYSKILPRYDKRNPHECVNSEM 2122
QY 824 LGTKOQVFOGTNSLGK--SSQSVOSIR-PPYRBAVSLDSPVSGSSPPYKNISAFPMPL 880
Db 2123 EKTKCKEFKLSNNLVNKGSSNNHSTIKVSPY-----LSQFOQDK 2162
QY 881 KQPMLGONPRMDQSENYGSSMGP--NRNVTYVOTPSSGDMGLPNKAGHMEPMNSN---936
Db 2163 OQVLVGTAKVLENIHVLGKEQASPKNVKMEICKTETFS-----VYKTIIEV 2211
QY 937 --SMGRGSGYNTSLRPAALGSGIPTILPILSNSNIPGARPYVLOOQOQMLQMRPGEIIPMGK 994
Db 2212 CSTYSKQSENY-----FETEAVEIAKAFMEDD-----ELTDSKLP---2246

QY 995 ANPYGAASNOIGSWPDG---MLSMEOVSHGTONRPLRNSLDLYGPPS---NLEGOS 1048
Db 2247 -----SHATHSLFTCEBNEEMVLSNSRICK-RGGEPLI-----LVGEPSIKRNL-----2289
QY 1049 DERALLDQHLTLSTNDATGLEEIDRALGPELVNOGALPEKQDAFOGEAAVYMDOKA 1108
Db 2290 -----LNEFDRI-----IENQKSLKASKSTPDG-----TIKDRLL 2320
QY 1109 GLYCOTYPAGCPMGOGFPHLOGSPFSFNSMMQNOGNEPPLQCMHPRAINMPRTYTPK 1168
Db 2321 FMH-----HV-----SLEPITCVPEFRTKEROEIONPNTAP- 2352
QY 1169 QLRMQLOQRLOGOFLNOSRQALEKMNPTAGGAAYVRPMQOPQOGFLNAQVARSRE 1228
Db 2353 -----GQEFLSKSHLHEHLETKSSNLAVSGHFFYQ-----VSATRNE 2391
QY 1229 LLSH-----HFRQRYAMMMQOQOQOQOQOQ 1255
Db 2392 KMRHLITTGRTKVFVPPFKTSHF--HRYEQCVRNINLEENRQK 2435

RESULT 6
US-09-421-124-44
Sequence 44, Application US/09421124
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
TITLE OF INVENTION: susceptibility gene and uses thereof.
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: NC 27622
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 3418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-421-124-44

Query Match 2.2%; Score 163; DB 5; Length 3418;
 Best Local Similarity 17.6%; Pred. No. 0.0067;
 Matches 251; Conservative 195; Mismatches 480; Indels 500; Gaps 60;

```

QY 7 NUDPLASDRKKRLPC-----DTPGGGLTCSGGRKRRREDSKYIEELALISAN 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1333 NIEFGSSSSKKDYCIHKDETDLLFTDOHNICLSGO-----FMKGNQIRED 1383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 56 LSDIDNFVVKPKCALKEVYRQIQKEOGKTIISNDDVOKADVSTGGGVIDKDSLGP 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1384 LSDLFLFLEV-----AKAGEACGNTSNKEQLATATEQNIKRFELSDT----- 1426
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 116 LLLQALDFLEVVRREANIVFVSENVTOYLOKQEDLVNTSVYINILHEEDRKOFELKLPK 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1427 -FFQFASKNISVAKE-----LEFNKIVFEDQKPEELHNFSLNSLSDIRKKNKDIISLY 1480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 176 STVNGVSWTNEPQROKSHTFNCRMLKTPHDLIEDINASPEKROYEMOCALSOPRAM 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1481 EETDIVK-----HKLKE-----SVP-- 1496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 236 MEEDGLQSCMICVARITTGER--TFPSNPE-----SEITRHDLGKVVINIDTNSL 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1497 -----VGTGQVLTFQOGPERDEKIKEPILGLFHTASGKVKVIAKESL 1539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 286 RSSMRPFEDIIIRCIQRFPSINDGOS-----WSQNRHYQEAYLNGHAETPVYRPSL 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1540 -----DKVKNL-----FDEKEQGTSEITSPSHOMAKTLKRYEACKD-----LEL 1578
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 ADGTI-VTAQFTSKLFRNPVTNDRH-----GFVSTHLEOREN-----GYR 377
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1579 ACETIEITTAAPCKEMQNSLNDKNLVSIETVVPKRLSDNLCROTENLTKISIFLAVK 1638
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 378 PNPVNVGOGIRPPMAGCNSSVVGMSMSPNOGLOMPSRAYGLADSTT----- 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1639 VHENVEKETAESP-ATCYTNQSPYVIEMSALAFYTSCKRKTSVGOTSLLEAKKMLREGI 1697
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 --GQ--MSGARYG-----SSNIAS-----LTPGQWSPSSYONNNTGLNMSSEP 467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1698 FDGQPERITADYVGNLYLENNSNSTIAENDKNHLEKODTSLNNSMSNSYSYHSEVY 1757
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 468 HGSPPLAQNQIMISPRNRGSPKIASHQFSPVAGVHSPMASSGMTGNISPS-----SS 521
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1758 NDSRLSKNKKLD-----SGIEPVLAKEVEDQKNTSESKVSIANSVKD 1796
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 522 SISLQATISEGVGISTLSTLSSGP-----KLDNSFNMMITOPS----- 560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1797 ANAYQGTAVNEDICVEEIVTSSSPCKNKNNAIKLISINSNFEVGPAPFIAAGKIRLCSH 1856
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 561 ----KVSNDKSPLEFYCDONPVSSMCQ-----SNSRDHLSDKES--KESVYEGAENORG 611
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1857 ETIKKVKDIDSFYSKVIETKENNENKSKICQTKIMAGCYEALDSEDLHNSLDNDE--- 1912
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 612 PLESGRHKLLQDLTCCSSDDRGHSSLTJNSPLDSCKESVSUTSPSGVSSSTSGVCS--ST 670
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1913 -CSMHSHTVFADI---QSEELIIGHONNMSGLE-----KVSISPCDVSLETSDICKCSI 1962
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 671 SNMHSLLQEKHRIHLKLLONGNSPAEYAKITPAQATGKDTSSITSCGCGNVYKQEBLSEK 730
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1963 GKLHRSV-----SSANMTCGIRSTASGKS-----VQVSQSLNANQVFESE 2002
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 731 KKEN--MALRLTLDRDBSDALSKELQOVGVDNKMSQ-----CTSSITP--SSSQ 779
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2003 IEDSRKQVFSKVLFKSNESHDLTREENTAITRPEHLISQKCFSYNVVNSSAFSGSTAS 2062
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 780 EKDPKIKETSEEGSGDDINDLDAITLGD-----LSSDPYNNISISNGSH 823
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2063 GKQVSILESSLHKVGVLEEDLITREHSLHSPTRQNVSKILPRVDRNRNEHCNVSEM 2122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 824 LGTKQOVFGTNSLGLK--SSQSVOSIR--PPYNRAVSLDSPVSVSSPPVKNIISAFPMPL 880
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2123 EKTCKEKFRLSNLNVVEGSSSENNHSIKVSPY-----LSQFOQOK 2162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 KQPLGNGPRMDSQENYSSMGPR--NRNVITVOTPSSGDGMLPNSKAGRMPEMNSN--- 936
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 2163 QQLVLGKTVSLVENIHVLKEDQASPKVNMKEIGKTEYSD-----VPATNIEV 2211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 937 --SMGRPGDVNTSLPRPALGSIPTLLPLRSNSIGARPVLOOQOQMLMRPGELPMGM 994
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2212 CSTYSKDSERNY-----FETFAVIAIAAFMDD-----ELNDSKLP----- 2246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 ANPYGAANASNOLGWPDPG--MLSMEOVSHGTQNRPLRLNSLDLVGPPS--NLEQOS 1048
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2247 -----SHATHSLFTCPREEVWVLSNRIGK--RGEPLI-----LVGEPSIKRL---- 2289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1049 DERALLDQHLTLSTMTDGTGLEIDRALGIPELVNOGALERKQAFQGEAAVMMQKA 1108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2290 -----LNEFDRI-----IENQERSLASKSTPDG---TKIDRL 2320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1109 GLYGQTYPAQGPPOGPFHLOQSPSPFNSMMNQMOQNFLOGMHRANIMRPTNPK 1168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2321 FMH-----HV-----SLEPTICVPRRTTKKEQEIIONPFTAP- 2352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1169 QLRMOLQRLQGOQFLNOSRQALELKMENPTAGAAVVRPMQPOQGLMOMVAQRSRE 1228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2353 -----GQERLSKSHLYEHLTLEKSSSNLAVSGHPYQ-----VSATRNE 2391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1229 LLSH-----HFRQQRVAMMQQOQQOQQOQQO 1255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2392 KMRHLITGPRPKVFPVPEPKTSHE--HRVEQCVBNIMLEENROKQ 2435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-08-945-567C-4
; Sequence 4, Application US/08945567C
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARRNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michael H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; FILE REFERENCE: 1038-745 MIS
; CURRENT APPLICATION NUMBER: US/08/945,567C
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1833
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
; US-08-945-567C-4

Query Match 2.2%; Score 161; DB 4; Length 1833;
Best Local Similarity 19.3%; Pred. No. 0.004;
Matches 241; Conservative 194; Mismatches 494; Indels 330; Gaps 64;

```

```

Db 692 TSANGTATN-----FNVSDEDALVNAKDIAENLTLAKEIHTTKGTADTALQT 742
Qy 227 FALSOPRAMEBEDLOSCMICVARITTEGERFPSPNEPFIIRHDSGVNVIDN--- 283
Db 743 FTVK---VDENNADD-----ANATVGOQN-ANNQVNTLTLKENGINTIDKNGTV 792
Qy 284 ----SLRSMRPGFEDIIIRCIORFSLNDG-----OSMSQRHYOEAYLNG 326
Db 793 TGCINTTSLGAKKS-----TLNDGSLINKPFGSEIOYGADGVKAKANNNG 841
Qy 327 HAETPVREFSLADGTIVTAQTSKLEFNPVYNDRHGFVSTHFLQRBONGY--RPNPNVPG 384
Db 842 VVGAGI-----DGT-----TRITREIGFTGT-----NSGLDKSKPHLSK 876
Qy 385 QGIRPPMAGCNSSVCGMSPNGLQMPSSRAYGLADPSTGTGMSGARIGSSNINSLTP 444
Db 877 DGI-----NAGGKKTINIIOGEL---AONSHDAVYGGKIYDLKTELEKRISS-7A 922
Qy 445 GPGMOSPSY-----ONNNYGLNMSSPHGSPGLAPNO-----ONIMISPRNRGSPKI 492
Db 923 KTAONSLHEFSVADBOGNNTV---SNPYSYDTSKTSVYITRAGENGITTKYKNKGVYR 979
Qy 493 ASHOFSPVAGVHSPMASSGNTGNHSESSSLALQAISEGVGTSLSTLSSPGPKLDNSP 552
Db 980 GIDQ---TKGLTTPKLTGVNNGKGIVIDSONGONTIT-GLSNTLANVTNDKG----- 1028
Qy 553 NMNITOPSKYSNODSKS-----PLGF-----YCDONPVESSM 584
Db 1029 SVFTTQGNLIKEDKTRAAIVDVLISAGFNLOGNEAVDFVSTYDTPVNDGNATYAKV 1088
Qy 585 CQNSRBDHLSDKESKESSVEGAEN--ORGPLESKHKILLOLTCSSDDRHS--SITNSP 641
Db 1089 TY-----DOTSKTSKYVYDVNVDITIEVADKLGVTTLTLTSTGICANFALSNOA 1140
Qy 642 LQSSCKESSVYTPSPGVSS--STSGVSSSTNMHGSLLQEKHRIHLKLLQNG----- 692
Db 1141 TGDALVKAQSIYVHLNLTLSDIOTAKGASQANNAGVYDADGKVIYDSTDNKYQAQND 1200
Qy 693 ----NSPAEVA--KITAOATGKOTSSITSCGDGVNVVQEOULSPKK-----ENNALRLYL 742
Db 1201 GYVDKTRKFAKDLVAOQTPD--GTLAONMNVKSVINKEOVANDANKOGIENEDNAFYKGL- 1258
Qy 743 DRDDPSDALSKELQPOVEGVNKMOSCTSTIPSSOEKPKI-----KTETSE-- 791
Db 1259 -EKAADNKTNAAYTV--GDLNAVAQ--TPTFTAGDGTGTTAKKIGELITLTKGOTDINKLT 1315
Qy 792 -----EG-----SGDLNLDAILGDJTSSD-----FYNNS-----ISSNGS 822
Db 1316 DNNIGYVAGTDEFTVKLAKDLTLNLSVNAAGTKIDKGVSFVDSGQAKANTPVLISANGL 1375
Qy 823 HLGTK--QOYFOGTSLGLKSSOSVOSIRPPYKRAVSLDSPVSGSPPKNISAPMLP 880
Db 1376 DLGKVIISVNGKGTJTDANVOOLEVR-----NLGLGNAQNDNADNOVNIADIK 1428
Qy 881 KQPLMGNPRMMDSOEYSSGMPNRNVTVTQTP-----SSGD-----MGLPN 924
Db 1429 KPNSSSSSKR--TVIKAGTVLGGKGNDEKATLGCIQYGVCKDGNANGLSNVM--VKT 1485
Qy 925 SKAGRMEPM--NSNSMGRPGDYNTSLPRPALGSIPTLPLRSNSIPGAR----- 972
Db 1486 OKGSKKALLATYNAAGOT--NYLTNNPAEALD-----RINE--QGIFFHVNONGOE 1534
Qy 973 PVLQOQOQMLQMRPE--IPMGANPYGOA-----ASQOLGSWPDGMLSM--EYV 1020
Db 1535 PVLQOQNGIDSSASGKHSVAIGFOAKADGEAAVAIGROTAGNOSIAIGMAATGDOST 1594
Qy 1021 SHGTORPLRNLSDLVGPSPNLGEGSDERALLDOLHTLNS--TDPATGEIDRALG 1077
Db 1595 AICTGN--VYAGKHSAGLIDPSTVKA-----DMSYSGNNNOFTDQYDQYFVGANN 1644
Qy 1078 IPELVNQGALPEPKODAFQOEAAVMMDOKAGLYGQYTPAGCPMPOGCF 1126
Db 1645 ITVETESNVALGSNSAISAGTHAGTQAKKSDGAGTTTGTGATGYKGF 1693

```

```

RESULT 8
US-08-945-567C-3
; Sequence 3, Application US/08945567C
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: KLEIN, Michel H.
; APPLICANT: CHONG, Pele
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; FILE REFERENCE: 1038-745 MIS
; CURRENT APPLICATION NUMBER: US/08/945,567C
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-08-945-567C-3

Query Match 2.2%; Score 161; DB 4; Length 1992;
Best Local Similarity 19.3%; Pred. No. 0.0044;
Matches 241; Conservative 194; Mismatches 494; Indels 320; Gaps 64;

Qy 64 VRPDKCAILKEVYRQIRQIEQKTSINDDVQKA---DVSSTGQGVIVKDSIGPLLLQA 120
Db 738 LSPTLPSLIDQSSRNI---ELGNTIODKDKSNAASINDLNTGFMU--KNNNPNIDFVS 791
Qy 121 LDGFLFVFNREANIVPVE---NVTQLOYKQEDLVNTSVYNIHEDRKDFLK--NLPR 175
Db 792 TYDIVEFANQNTATYVHTFANKTSKYV--DVAVDDTTHLGTDDNKKLGKTKTKLNK 850
Qy 176 STYNGVSWTNEPQOKSHFNCRM---LMKTPHLLIEDINA--SPENRQRYET---MOC 226
Db 851 TSANGTATN-----FNVSDEDALVNAKDIAENLTLAKEIHTTKGTADTALQT 901
Qy 227 FALSOPRAMEBEDLOSCMICVARITTEGERFPSPNEPFIIRHDSGVNVIDN--- 283
Db 902 FTVK---VDENNADD-----ANATVGOQN-ANNQVNTLTLKENGINTIDKNGTV 951
Qy 284 ----SLRSMRPGFEDIIIRCIORFSLNDG-----OSMSQRHYOEAYLNG 326
Db 952 TGCINTTSLGAKKS-----TLNDGSLINKPFGSEIOYGADGVKAKANNNG 1000
Qy 327 HAETPVREFSLADGTIVTAQTSKLEFNPVYNDRHGFVSTHFLQRBONGY--RPNPNVPG 384
Db 1001 VVGAGI-----DGT-----TRITREIGFTGT-----NSGLDKSKPHLSK 1035
Qy 385 QGIRPPMAGCNSSVCGMSPNGLQMPSSRAYGLADPSTGTGMSGARIGSSNINSLTP 444
Db 1036 DGI-----NAGGKKTINIIOGEL---AONSHDAVYGGKIYDLKTELEKRISS-7A 1081
Qy 445 GPGMOSPSY-----ONNNYGLNMSSPHGSPGLAPNO-----ONIMISPRNRGSPKI 492
Db 1082 KTAONSLHEFSVADBOGNNTV---SNPYSYDTSKTSVYITRAGENGITTKYKNKGVYR 1138
Qy 493 ASHOFSPVAGVHSPMASSGNTGNHSESSSLALQAISEGVGTSLSTLSSPGPKLDNSP 552
Db 1139 GIDQ---TKGLTTPKLTGVNNGKGIVIDSONGONTIT-GLSNTLANVTNDKG----- 1187

```



```

APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-421-124-16

```

```

Query Match      2.18; Score 158.5; DB 5; Length 2329;
Best Local Similarity 17.68; Pred. No. 0.0078;
Matches 250; Conservative 195; Mismatches 482; Indels 495; Gaps 59;

```

```

QY 7 NLPLASDSRRKRLPC-----DTPGQGLTSGEKKRRRQESKYIELAELISAN 55
DB 1052 NLEFDSDSSKNDVTCHKRETDLLFTDQHNICKLSGQ-----FMKEGTQIQED 1102
QY 56 LSDIDMENVKPDCAILKETVROIKEQKTSNDVDYQKADVSTGGVIGDKDSLGP 115
DB 1103 LSDLTFLFV-----AKQOEACHGNTSNKEQLTATKTEQNIKDFETSDT----- 1145
QY 116 LLLQALDGLFVYNNREANIVFVSENVTOYLOKQEDLVNVSVINIILHEEDRKDFLKLPLK 175
DB 1146 -FPTQASGKNISAVAKES-----FNKIYVFPDQKPEELHNFSLNSELSDIRKKNMDILASY 1199
QY 176 STVNGVSWTNEPQOKSHFTFCNMLKTPHDLIEDIVASPEMGRVETMGCFLASOPRAM 235
DB 1200 EETDIYK-----HKILKE-----SVF----- 1215
QY 236 MEEGEDLSQCMICVARRITTEGR--TFPSNPE-----SFTIRHDLGKVVNIIDTNSL 285
DB 1216 -----VGTGNQLVTFQGGPDERDEKIKEPTLLGFTHTASGKVKYIAKESL 1258
QY 286 RSSMRPGFEDITIRRCIOREFSLNDGOS-----WSQKRHYQDAVYNGHAETPPVYRFSL 337
DB 1259 -----DKVKNL--FDEKEGEGTSEITSFSHOWAKTLKYREACKD-----LEFL 1297
QY 338 ADGTI--VTAOTKSKLFENPVTNDRH-----GFSVTHPIFORBON-----GYR 377
DB 1298 ACETIETITAPKCKEMONSINNDKNLYSITVPPKILLSNLCROTENLTKSISIFLKV 1357
QY 378 PNPNPVGQGIKPPMAGCNSSVGMSMSPNOGLMPPSSRAYGLADPSTT----- 425
DB 1358 VHENVEKETAKSP-ATCYTNQSPSYIENSALAFYITSCSRKTSYSGQSLLEAKMKWLEGI 1416
QY 426 --GQ--MSARVGG-----SSNIAS-----LTPGPKQSPSSSQNNNYGLNMSPP 467
DB 1417 FDOQPERINADYVGNLYENNSNSTAENDKNHLSKODTYLSNMSMSYSYHSDDEVY 1476
QY 468 HGSFGLAPNQNIMISPRNGSPKIASHQSPVAGVSHSPMASSGNTNHSFS-----SS 521
DB 1477 NDSGYLSKNLID-----SGIEPVAKNVEDQNTSFSKVIYSVWKD 1515
QY 522 SLSSALQAISSGVGTSLISTLSSPGKILNPNMNTIQPSKVSNDKSPSLGFCYCDQNPVE 581
DB 1516 ANNAVQTVNEDICVEELVYTSSSPCKNKNAIKILISN-----SNNEFVGPAFRIASGKI- 1570

```

```

QY 582 SSMQSN-----RDHLSDEKSESSVEGAENQRPLESKGHKKLLQLLTCSDDRGHSS 636
DB 1571 --VCVSHETTKKVKNDIFTDSPFSKYK--ENNENKSKICOTIMAGCYEALD--DSEDLIHN 1626
QY 637 LTNSPLDSSCKESSVSVTSPPSGVSSSTSGGVSSNNHSGSLQEKHRIILKLLQNGNSPA 696
DB 1627 LDN-----DECSHSHKVFAD-----IQSEILQIN--QNMSSGLE 1659
QY 697 EVAKITQATGKDTSSITSCGDDNV-----VKQELSPKKKEN----- 734
DB 1660 KVSKITSPCDVLSLETSIDICKSIGIKLHKSVSANTCGIFSTAGSKSVQVSDASIQNAROY 1719
QY 735 -----NALRYLLDRDPDALSKELOPQVEGVNKKSQ-----CTSSITP--SS 777
DB 1720 SEIEDSTKQVSKVLEKSNHSDQLTREENTAIPTPHLISQKGFSYNVNVSASFSGFST 1779
QY 778 SOEKDPKIKTETSSESGDDNLDAIIGD-----LTSSDFYNNSSISSNG 821
DB 1780 ASGQVSLSSSLKVKVGVLEEDLITRHSLSHYSPTRQNVSKILPRVDKRRPHECVNS 1839
QY 822 SHUGTKQVFOGTNSIGLK--SSQSVOSIR--PPYNRAVSLDSPVSGVSPPVKNISAFPM 878
DB 1840 EMEKTSKEFEKLSNNLVEGGSSSENNHSIKVSPY-----LSQFOQ 1879
QY 879 LPKQPMLGPNPRMMDQENTGSSMGP--NNNVYVOTRPPSSGDMGLPNSKAGRMEMNSN- 936
DB 1880 DKQQLVIGTIVSLVENIHVLGKEQASPKNVKMEKTKETESD-----VPVYNTI 1928
QY 937 -----SMRPGGDVNTSLPRPALGSIPTLPLRSNISGARPVLOQOQMLMRGGEIIPMG 992
DB 1929 EVCSITYSKDSENT-----FETPAVIAIAFMEDD-----ELDSKIP-- 1965
QY 993 MGANPYQAAAQNLGSGWPDG--MLSMEOVSHQTONRPLRLNSLDLVGPPS--NLEG 1046
DB 1966 -----SHATSHLFCPENEMEWVLSNRIGK--RGEPLI-----LVGEPSIKRNL-- 2008
QY 1047 QSDERALLDQHLITLSTWDATGLEIDRALGITEPLVNOGALPEKQDAFOGQEAAYVAMQ 1106
DB 2009 -----LNEFDRI-----IENQESLAKSKSTPQG-----TIKDR 2037
QY 1107 KAGLYGOTYPAQGPBMGGFHLQGOQSPFSNMNQNMQGNFLOGHPRANIMRPPTNT 1166
DB 2038 RLRYVH-----HV-----SLEPTICVFPRTTKEQEOINPFT 2070
QY 1167 PKOLRMQLQRLQGOQFLNOSROALELKMENPTAGAAVNRPMQOQGLFNAQVAAORS 1226
DB 2071 P-----GQEFLSKSHLYEHLTLEKSSSNLAVSGHPFYQ--VSGKNGKM----- 2112
QY 1227 RELI-----SHHFRQQRVAMMAMQOQOQOQOQ 1252
DB 2113 RKLITTGRTKRVVPEPKTSFHRVQCVRNINLEGNROKQ 2154

```

```

RESULT 11
US-09-421-124-16
Sequence 16, Application US/09421124
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSER: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: NC 27622

```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-421-124-16

Query Match      2.1%; Score 158.5; DB 5; Length 2329;
Best Local Similarity 17.6%; Pred. No. 0.0078;
Matches 250; Conservative 195; Mismatches 482; Indels 495; Gaps 59;

QY 7 NLDPLASDSRKKRLPC-----DPPGGGLTCSGKRRRREGSKYIELAELISAN 55
DB 1032 NLEFGSDSSKNDYVCIHKEDEFLFTDOHNICLKLSGQ-----FMKGGNTQIKED 1102
QY 56 LSDIDNFNKPDKALIKETVROIKRQKGTISNDDVDQKADVSTGQGYIDKDSLGP 115
DB 1103 LSDIDFLFLEV-----AKAQCHGNTSNKQLATKTEQNIKQEFERSDT----- 1145
QY 116 LLLQALDGLFVYNREANIVFVSENVTOYLQKQEDLVNTSYNNILHEDRDKDFLKNLPK 175
DB 1146 -FFQVAGSGNISVAKES-----FNKIVNFDDQKPELHNFSLSNSELHSDIRKKNKDILSY 1199
QY 176 STVNGVSWTNEPQROKSHTFNCRMLKTPHDLIEDINASPENKRORETWQCFALSOGRAM 235
DB 1200 EETDIYK-----HKILKE-----SYV--- 1215
QY 236 MEEGEDLOSMTICVARRTTGER--TFPSNPE-----SEITRHDSGKVVNTIDTNSL 285
DB 1216 -----VGTGNQLVTFOGQPERDEKIKEPFLILGFTAGSKKVKIAXESL 1258
QY 266 RSSMRPGEFDIRCTQRFFSLDNGOS-----WSQKRHYQEAVALNGHAETPVYRRSL 337
DB 1239 -----DKYKNL--FDEKEOGTSEITSPSHQWAKTLKRYREACKD-----LEL 1297
QY 338 ADGTI--VTAQTSKLFERNVVTNDRH-----GFVSTHLEOREN-----GYR 377
DB 1298 ACETIEITIAAPCKEKQNSLNDKMLVSETIVPPKLSLDNLCROTENLKTSKSIFLKVK 1357
QY 378 PNPNVGQIRPPMAGCNSSVGMSPNOGLQMPSSRAYGLADSTT----- 425
DB 1358 VHENVEKEIKAKSP-ATCYTNOSPYSVIENSALAFYTCGRKTSVSQTSILLEAKKMLREGI 1416
QY 426 --GQ---MSGARYG-----SSNIAS-----LTPGFGMOSPSSSYONNNGLMMSSPP 467
DB 1417 FDGQPERINTADYVGNLYLENNNSNTIAENDKNHLSKQDTYLSNMSMSYSYHSDEVY 1476
QY 468 HGSPLAPNQOQIMISPRNRGSPKIASHOFSPVAGVHSPMASSGNTGNHSFS-----SS 521
```

```
DB 1477 NDSGLTSKKNKD-----SGLPEVLKAVEQKNTSEKVIANSVKD 1515
QY 522 SLSSALQATSEGVGTSLSLSTSSPGPKLDNSPMNMTTOPSKVSNODSKSLPGFYCDQNPVE 581
DB 1516 ANAYQTVMEDICVELYVSSSPCKNKNKNAITLSTSN-----SNNEVGPAPRIASGRI- 1570
QY 582 SSMCOSNS-----RDLSDRKESSESSVEGAENQRPLESKGHKLLQTLTSSDDRGHSS 636
DB 1571 --VCVSHETIKRKVDIFDTSPFSKVIK-ENNEKSKRICOTIKWAGCYEALD-DSEIDLINS 1626
QY 637 LFNPSPLDSSCKESSVSVTSPSGVSSSTSGGVSTSNMHSGLQEKHRIHLKLONGNPA 696
DB 1627 LDN-----DECSTHSHKVPFD-----IQSEELIQHN--QMSGLE 1659
QY 697 EVAKITTAQATGKDTSSITSCGDGNV-----VKOEOLSPKKEN----- 734
DB 1660 KYSKISPCOVSLSTEDICSGIKLHKSVSANNTGIFSTAGSKSVQVSDASLOAARQVF 1719
QY 735 -----NALLRYLLDRDPSDALSKELQPOVEGVNKNMSO-----CTISSTP--SS 777
DB 1720 SEIEDSTKQVFESKVLFSKSNHSDQLTREENTAIRPEHLISQKGFSYNVVNSAFSGCFST 1779
QY 778 SQEKDPKIKETSEBEGSGDLNLDALILGD-----LTSSDPFNSSISNG 821
DB 1780 ASGKQVSLLESSLHKVKGYLEFDFLRTHEHSLHYTSPSRQNVSKILPRYDKRNPHECVNS 1839
QY 822 SHLGTQQVFOGNTSLGLK--SSQSVQSIK-PPYNRAVSLSPSVSGSPRYKNISAPFM 878
DB 1840 EMKETSCKREFKLSNNLNVGSGSENNHSIKVSPY-----LSQFQQ 1879
QY 879 LPKQPLGPNPMMDSQENYSGSKGP--NRNVTYVQTSPSSGCMGLPNSKAGMEPMNSN- 936
DB 1880 DKQQLVLTGKSVLSENIHVLGKEQASPKNVKMEIKETFEFSD-----YVAKTNI 1928
QY 937 ---SMGRPGGDYNTSLPRPALGGSIPTLPLNSNIPGARPYLQOQOQMLQRPPEIIMG 992
DB 1929 EVCSTYSKSENY-----FTEAVEILKAKMEDD-----ELTDSKLP- 1965
QY 993 MGANPYGGAASNOLGMPDG---MLSMEOVSHGTQNRPLRNSLDDLVPSPS---NLEG 1046
DB 1966 -----SHATHSILFTCPENEMEVLNSNRICK-RRGEPLI-----LVGEFSIRNL-- 2008
QY 1047 QSDERALLDQLHTLNSNDATGLEETDRALGIPELVNGQALERPQDAFQCGEAAVMMQDQ 1106
DB 2009 -----LNEDRF-----IENEKSLKASKSTPDG---TIRDR 2037
QY 1107 KAGLVGQFYPAQGPWMGGFHLQGGSPSPNSMMNMOGQNPPLGMPHRANIMRPNTT 1166
DB 2038 RLFFVH-----HY-----SLEPTTCVFPFTTTEREQTQNPFTA 2070
QY 1167 PKQLRMQLOQRLOGOQFLNDSROALELKMENPTAGAAVMRPMQPOQFLNAQVAAQS 1226
DB 2071 P-----GQEFLSKSHLYEHLTLEKSSMLVAAGHPFYQ-VSGNNKNGK----- 2112
QY 1227 RELI-----SHHFRQQRVAMM0Q000000 1252
DB 2113 RKLITTGPRTKVVPFPKTSHFHREVCVRNINLEGNRQKQ 2154

RESULT 12
US-60-248-505-927
; Sequence 927, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 927
 LENGTH: 2128
 TYPE: PRT
 ORGANISM: Human
 US-60-248-505-927

Query Match 2.1%; Score 155; DB 6; Length 2128;
 Best Local Similarity 17.7%; Pred. No. 0.012;
 Matches 195; Conservative 133; Mismatches 419; Indels 352; Gaps 40;

```

OY 466 PPHSGPLAP-----NQNMITS 483
DB 457 PPGSGEAPLAGAASLQDEVREAEAGRILTERENRELGLLOVLOGQPGQHLLER 516
OY 484 PR-NRGSPIASHQSPVAVGNSP--MASSGNTGNHSSSSLSALQAISEGVSTLLST 540
DB 517 PREDPVLPLEEAPQTPVAFDHSPOGLVQKARDGCPQALDAPALDVLASAECPQAP 576
OY 541 LSSP-----GPKLNSPMNITQPSKVSMDS--KSPLGFTYCOQNPV 580
DB 577 DSDPOAESPLQAAAMPQASDMSPOEGSPVETQESPEKAGRRSSLQSPASVAPPQPG 636
OY 581 ESSMCSNRDHLSDKESSESV-----EGAENORGPL-----ESKHKKLLQ 623
DB 637 TKIQAPQLLGEGEGREAPQGLVPEAMGLRQGPPEHKGPSPSSVQLEQEGCPNQGDL 656
OY 624 LITCSDDDRGHSS-----LTNSPLDSSCKESSVYTPSPGVSSTSGVSTSNMH 674
DB 697 LATGQAEAREHQRLGRTVDPAMQPKQKSEGALFVQWEGPIPESLASGVAEQEALR 756
OY 675 GSLLQEKHRI-----LHKLLQNGSPAEVAKITQATCKD----- 710
DB 757 EEVAQLRRRAEALGDELEQARLEAONTAEALSKLEQARRAEAEHREAEQAQMEQA 816
OY 711 ---SSTSCGDGNVVKQEOLESPKKENNALRLRLDRDPS--DALSKELQPOVEGDNKA 766
DB 817 RLRENVAAAG-----QLEESASQREALVEAAAGRREROMREBSRLRAQGEAAEER 869
OY 767 SOCTSTIPSSOEKDPKTKTETSEEGSDLDLDAIILDLTSSDFYNNSSISSGHL-- 824
DB 870 ---QLESEGRQHLLEAEERERERK-----FALQALEKAVVRGKELGARLHLQR 916
OY 825 GTRQOAFQGTNSGLKSSQSVOSIRPPYNAVSLDSPVSGSSPPYKNISAFMLKQPM 864
DB 917 ELEQAALEQOELFREKESQ-----HORYQGLEO----- 944
OY 885 LGSNPRMDSQENYSSMGSGPNRNTVTQTPSSGDMGLPNSKAGRMEMNNSMGRPGD 944
DB 945 -----RLE----- 947
OY 945 YNTSLPRPAGGSIPTLPLRNSIPGARVYLQOQOQMLQNPGEITPMGMCANFYGOAAS 1004
DB 948 --AELQAAATSKREALMELKTR-----LOLEELFOLRQ--DAGLGP---KKRAE 992
OY 1005 NOLGSPMDMLSMEOYSHG--TQNRPLRLNSLDDIVGPPSNLEGQDERALLDQL---HT 1059
DB 993 POLVETQNRLLFEVERSNMALVAEKALQGLQHLGQLSLQGRQOELLLOSORAQESH 1052
OY 1060 LLSNTDTATGLE---EIDRALGIPELVNOGQALEPKQDAFOGQEAIVMMQDQAGLQGY 1115
DB 1053 SRLQAKSVLEIQGGLHNRKLEV--LEEFYBARAQSEEFRRGOQOALLBDHKA---QLD 1106
OY 1116 PQAQPPMOC-----GFH--LOGQSFENS-----MNO 1141
DB 1107 RQGEALTEGILVHRDKLKNMRALETAHRELQGRHQLQORASVAEQVLAERERLMD 1166
OY 1142 MNQGNFP-----LOGMHRANIMRPTNPK-----OLRMDL 1174
DB 1167 GHRQRLLEELRLQSEHDAQMLLAELSKERELOGERGEILGRILAELEPAQLEMOS 1226
OY 1175 QORLQGOQFLNRSQALELKMENPTAGAAVNRPMQPOQGLFMAQVARSRELL----- 1230
  
```

```

DB 1227 QOLRESNOQLLSACRLTQCCELLTQLRSQAE---ENROLAEVQALSHRENRELSL 1283
OY 1231 ---SHFROQRV-----AMMQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 1280
DB 1284 ESRDHLHREQREYLDQMLMRREKQLYEKIMQYRLVEPEGLAGQSGSPSPAPRRRQ 1343
OY 1281 PSM---DGLLAGPTMPQAPQOQFPY---QPNYGMGOQPDPAF--GRVSSPPNMAWSSRM 1333
DB 1344 SSLCLDETLGAGQRRKL--SRFPVGRSSSESPPGDTPROFRQHRPGLGAFVSHSKGP 1402
OY 1334 -----SONPMQHPQOAS 1346
DB 1403 GVGWENSAETLOEHENDAN 1421
  
```

```

RESULT 13
US-60-248-505-690
; Sequence 690, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c10009318
; CURRENT APPLICATION NUMBER: US/60/248,505
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 690
; LENGTH: 884
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-690
  
```

Query Match 2.1%; Score 151.5; DB 6; Length 884;
 Best Local Similarity 19.9%; Pred. No. 0.0061;
 Matches 177; Conservative 112; Mismatches 386; Indels 215; Gaps 39;

```

OY 660 SSSTSGGVSTSNMHSGLQEKHR-----LHKLLQNGN----- 693
DB 43 ASDSDSGATLTGRELKQLQGEFGDFQPCVLHAAVEKNSNLINDSGTISFDEVLATF 102
OY 694 -----SPAEVAKITQATGK--DTSSTSCGDGNVVKQEOLESPKKENNAL 738
DB 103 NLNLICYLDIKSLSELRLQVTRPEKEKLDVDVQATTGG-----QWTVGTSTPOEKRML 158
OY 739 RYLLDRDDPSD--ALSKELQPOVEG-----VDKMSQCTSTIPSSOEKDPKIK---T 787
DB 159 -----PSGMASSQQLIPBESGAVGNRVDPRWRAKTHNPFGEASEHNDFPKHLEGD 210
OY 788 ETSSESGDLDNLDAIILDLTSSDFYNNSS--ISSNSHLQTKQOVFOGTNSLGLKSSQSV 845
DB 211 EGOGEVAQDIQIOTJEDNEGQKTKPKMAGSKKTSPTERRGQDKEISQEGDEPA--REG 268
OY 846 QSTRPPYNAVSLDSPVSGSSPPYKNISAFMLPKQPMILGNP-----KMSQENYSGS 901
DB 269 SKTRDQFGEQ--EGNLAQOSSPP--KEATQRPCEDETQKSTDSKDYCRMFDOE----- 319
OY 902 MGPNNRNTVTQTPSSGDMGLPNSKAGRMEMNNSMGRPGDYNLSPRALGSGSIPTL 961
DB 320 ---PGKDAQOT-----PATKMLGER--EYGGTSETOKECECTKDL 356
OY 962 PLRNSIPGARVYLQOQOQMLQNPGEITPMGMCANFYGOAASNOLGSPMDMLSMEOYS 1021
DB 357 PVQYGRNS-----ETSDMR--DEKERRRQPEAHGTAQGERBRKW---LVLETOG 403
OY 1022 HGQNRPL--LRNSLDDIVGPPSNLEGQDERALLDQLHTLSNTDTATGLE----- 1070
DB 404 QDKQYQELGLSKSKDAERG--SETQYLSSEGG--DQTHPELEGTAVSSEAEHTKEGTA 459
OY 1071 -----EIDRALGIPELVNOGQALEPKQDAFOGQEAQV-----MDQKAGLYGOT 1114
  
```


RESULT 15
US-60-248-505-714
; Sequence 714, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 714
; LENGTH: 2000
; TYPE: PRT
; ORGANISM: Human
; US-60-248-505-714

Query Match 2.08; Score 147; DB 6; Length 2000;
Best local similarity 18.5%; Pred. No. 0.034;
Matches 168; Conservative 130; Mismatches 342; Indels 268; Gaps 35;

QY 425 TQOMSGARYGSSNIA-SLTGPGKQSP-----SSY-----QNNNYG----- 460
DB 3 SSRSGIYGVGEOSTYVQDSATPSPPPPQPTERTPTSAIYRLRQVSYPTAECCOHLGI 62
QY 461 LNMSSPPHGS--PGLAPNOQIMISPRNRGPKIASHOFSVAGVHSPMASSGNTGNHSF 518
DB 63 LCICSRGCGTRVPSLLPQHDSVPPASARATTPSSSFVOTEP---FHPEQASSTQOQDGL 119
QY 519 SS--SSLSALQAISEG---VGTSLSTLSPPGPKLDNSPNNNITO--PSKVSMDPSPL 571
DB 120 LNRPSAFSTVQSTAGNTLRNLISLGPTRRSILGPLSHPSRYHREIAPGLGSEWTRIVL 179
QY 572 GFYCDQNVESMCOQSNRDLSDKESKESVEGAEENGRPLESGKHKLQQL----- 625
DB 180 SL-----NSRSE-----AESMPPTSAVSLSLVLRQDECG 212
QY 626 -----TSSSDRGHSITNSPLDSSCKESSVTSFGVSSSTGCVSSTSNMHSLLQ 679
DB 213 SQASVYITATGCRGPPA-----SGLATESDGGNGSSONNSGSTRH 252
QY 680 E-----KHRLHKLQ--NGNSPAEYAKITAOATGKDTSSITSCGDGNNYKQEBL 727
DB 253 ELQCDLRRFLELDRLQELDLSGEAPQ-----TQAOEMLNNNIESPPGSH-----QP 302
QY 728 SPKKENNAL-----RYLLDRDPSDALSKELQPOVEGVNKKMSOCTSS--TI 774
DB 303 TPHSSENNNSLRGHLNCRACHNLLTFNNNDTLRWERTTPNYSGEASSWQVPSSFESEV 362
QY 775 PSSSOEKDPKIKTETSEEGSGDLNLDALIGDLTSSDFYNNIS--SNGSHLGTRQOVPFG 833
DB 363 PSSGSQPLPLERTGQTPSSSRL-----ELSSDOERTVGVAFNOETGHEMERITYQ 414
QY 834 TNSLGLKSSQGVQSIIRPPYR--AVSLDSPVYSGSSPPVKNISAFPMILPKOPMLGNPR 890
DB 415 SSRGTYVQEQALHODMPESSEEDSLRLSPAAYVAQRMIOYLSRDSIRQSMRYQONR 474
QY 891 MMDQENYGSMMGPNRNVITYQTTPSSGDWGLPNSKAGRMEMPNSNSMGREGDYNTSLP 950
DB 475 LRSSTSSSSSDNOGVSVEGTLDEFEDFNG-----DRSRHAPRNARMS 519
QY 951 RPALGGSIPTLPLRSNSIPGARPLYLOOQOQMLQMRPELIPMGANPYGQAASNQLGSM 1010
DB 520 APSLGRFPYPRFLPEYLYPA-GIHERGQ-----PG----- 550
QY 1011 PDGMLSMQVSHGTQNRPLRLNSLDDLVGPPSNLEGOQSDERALDQTLTLSTNDATGIE 1070
DB 551 -----LATHSSVRYVLGAVIGD-----GQS-----AVNIANTYLYLQMMDFTKFD 589

QY 1071 EIDRALGIPELVNOQGALEPKQDAFQGBAAV-IMDQKAGLY-----GQTYPAQG 1119
DB 590 -----LPEISN-----ASVNVLYQCKIKYINDASCDISADGQLAIFI 626
QY 1120 PPMQGFHLQGSQSPSF-----NSMNNOMNOQGNFPLQGHHPANIMR 1161
DB 627 PSSQKGFDEGLAVYSLAPHNLGEMLYTKRGPNAISVSLSPMGRYVMVGLASRRILLH 686
QY 1162 PRTNTPKQLRMQLOQRLOGQOFLNDSROALEIKMENPTAGGAA-----VMRPMQ-P 1212
DB 687 PSTEHMVAQVFRLOQAHGGETSMREGVAVAKKDVIHMPKHELADKNVDNLHVKKAMQSLK 746
QY 1213 QQGFLLNAQ 1220
DB 747 SRGYVKEQ 754

Search completed: April 19, 2001, 16:54:07
Job time: 491 sec

OY 64 SLQEKHRIHLKLLQNSPAEVAKITQAOTGKDTSSITSCGDNVVKQOLSPKKENN 123
Db 67 S-----SSSSDSESEAFETKREBSKSSSSSSSDSEEBEKEETKKEES 113

RESULT 2
US-08-320-559-30
; Sequence 30, Application US/08320559
; Patent No. 5631135
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Cnaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods for
; TITLE OF INVENTION: Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
; TITLE OF INVENTION: All-1 Region
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5631135rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/320,559
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/062,443
; FILING DATE: 14 MAY 1993
; PRIOR APPLICATION DATA: US/07/971,094
; APPLICATION NUMBER: US/07/971,094
; FILING DATE: 30-OCT-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,830
; FILING DATE: 27-MAY-92
; PRIOR APPLICATION DATA: US/07/805,093
; APPLICATION NUMBER: US/07/805,093
; FILING DATE: 11-DEC-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-0855
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-320-559-30

Query Match 11.7%; Score 82; DB 1; Length 568;
Best Local Similarity 24.7%; Pred. No. 1;
Matches 42; Conservative 20; Mismatches 56; Indels 52; Gaps 5;
OY 7 KLLQL-----LTGSDRGHSLTNSPLDSCKRESVSVSPSGVSSSTGCVST- 58
Db 133 KLLKAGGDNRSIRHTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSTS 192
OY 59 -----SNMHSILOEKHRIHLKLLQNS-----NSPAEVAKITQAOT 94
Db 193 FSKPHKLMKREKPKSDSREHKSASFKEPRSDHKKSSKESKKPKKEKPLKEKIVPKMA 252

OY 95 GKDT-----SITSCGDN-----VVKQOLSPKKENNA 124
Db 253 FKEKPKSKREPKPDSNLTITSGDDKKAPKRPPISSSELSAKKRRKS 302

RESULT 3
US-08-545-860D-30
; Sequence 30, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Cnaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; STREET: No. 6040140rls
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994
; PRIOR APPLICATION DATA: PCT/US92/10930
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA: US 07/971,094
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA: US 07/805,093
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-545-860D-30

Query Match	11.1%	Score 78	DB 3	Length 506
Best Local Similarity	35.4%	Pred. No. 2.3		
Matches 23	Conservative 10	Mismatches 28	Indels 4	Gaps 1

RESULT 7
US-09-273-565-19
: Sequence 19, Application US/09273565A

```

? TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
? TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
? FILE REFERENCE: 0-53599
? CURRENT APPLICATION NUMBER: US/09/273,565A
? CURRENT FILING DATE: 1999-03-22
? EARLIER APPLICATION NUMBER: 09/055,699
? EARLIER FILING DATE: 1998-04-07
? EARLIER APPLICATION NUMBER: 08/920,170
? EARLIER FILING DATE: 1997-03-19
? EARLIER APPLICATION NUMBER: JP 63410/1996
? EARLIER FILING DATE: 1996-03-19
? EARLIER APPLICATION NUMBER: JP 69163/1997
? EARLIER FILING DATE: 1997-03-05
? NUMBER OF SEQ ID NOS: 95
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 19
? LENGTH: 506
? TYPE: PR1
? ORGANISM: Homo sapiens
? SS-09-273-565-19

```

Query Match	11.1%;	Score 78;	DB 4;	Length 506;
Best Local Similarity	35.4%;	Pred. NO.	2.3;	

RESULT 8
US-08-755-587-189

```

: RESULT 8
: US-08-755-587-189
: Sequence 189, Application US/087555587
: Patent No. 6045997
:
: GENERAL INFORMATION:
: APPLICANT: Futreal, Phillip A
: APPLICANT: Wooster, Richard F
: APPLICANT: Ashworth, Alan
: APPLICANT: Stratton, Michael R
:
: TITLE OF INVENTION: Materials and methods relating to the
: TITLE OF INVENTION: identification and sequencing of the BRCA2 cancer
: TITLE OF INVENTION: susceptibility gene and uses thereof.
:
: NUMBER OF SEQUENCES: 222
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Bell Seltzer Park & Gibson
: STEETER: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
: City: Raleigh
: STATE: NC
:
: COUNTRY: USA
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/755,587
: FILING DATE: 25-NOV-1996
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9523959.6
: FILING DATE: 23-NOV-1995
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9525555.0
: FILING DATE: 14-DEC-1995
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9617961.9
: FILING DATE: 28-AUG-1996
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Kenneth D Sibley
: REGISTRATION NUMBER: 31,665
: REFERENCE/DOCKET NUMBER: 5405-135
: INFORMATION FOR SEQ. ID NO: 189:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1589 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
: IS-08-755-587-189

```

Query Match	11.1%	Score 78	DB 3	Length 1589
Best Local Similarity	24.1%	Pred. No. 12		
Matches 40	Conservative 20	Mismatches 62	Indels 44	Gaps 6

```

QY      2  ESKHKKLLDLCSSDDRRHSLL--TSPDLSSCKESSVSUTSPGAGSSNPGCVSSTS  59
Db      1319  EMDGAKQLSSMLEGNEKPHSKRENSVHSTQGVLSLPRPLGANNSSVFSGSESTAG  1378
QY      60  NMHSSLQEHRLILHKL-----LQNGNSPAEYAKI-----  89
Db      1379  ---GKLVTSESALHKYKGLKEEFDLIRTEHLQHSPIREDYSKILPRQCAEIRTPRY  1455
QY      90  -TAQATGKDTSSI-----TSCGDGVNKKOE---QLSPKKKENALL  126

```



```

CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/667, 276a
FILING DATE: 11-MAR-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 377.5351P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-667-276A-6

```

	Query Match	23.08;	Score 77.5;	DB 1;	Length 254;
	Best Local Similarity	11.49;	Pred. No.1;		
	Matches	26;	Conservative	20;	Mismatches 64; Indels 1; Gaps
OY	14 TCSSDDRRHSLTSPDPSCKESSEVSTPSPGVSSSGSVSTSMNHGGLQEKIRIL	73			
	::: :				
Db	126 TSSAPSSSAAPTSSAASSSEAKSSSAAAPSSSEKSSAAPSSEAKSSSAAPSSSEAK	18-			
	::: :: :				
OY	74 HKLLONGNSPAEVAKITQAATGTGDTSSITSCGDGNVVKQEOLSPKKENNA	124			
	::: :: :				
	186 SSSAAPSSTEAKITSAPSSSTGGATTSALSTTDIOIQTAKVVS-EQENGCA	235			
	::: :: :				

RESULT 12
 US-08-776-265-5
 Sequence 5, Application US/08776265
 Patent No. 6001631
 GENERAL INFORMATION:
 APPLICANT: BLANCHE, Francis
 APPLICANT: CAMERON, Beatrice
 APPLICANT: CROUZET, Joel
 APPLICANT: FAMECHON, Alain
 APPLICANT: FERRERO, Lucia
 TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
 TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farbow, Garrett &
 ADDRESSEE: Dunner, L.L.P.
 STREET: 1300 I. Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/776,265
 FILING DATE: 24-JAN-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

```

:      NAME: Einaudi, Carol P.
:      REGISTRATION NUMBER: 32,220
:      REFERENCE/DOCKET NUMBER: 03806.0394-00000
:      TELECOMMUNICATION INFORMATION:
:      TELEPHONE: (202) 408-4000
:      TELEFAX: (202) 408-4444
:      INFORMATION FOR SEQ ID NO: 5:
:      SEQUENCE CHARACTERISTICS:
:      LENGTH: 663 amino acids
:      TYPE: amino acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: peptide
:
US-08-776-265-5

```

Query Match	11.08	Score 77	DB 3	Length 663
Best Local Similarity	22.99	Pred. No. 4.4	48	Indels 64
Matches 40	Conservative 23	Mismatches		Gaps
QY	16	SSDRGHSLNSPLDSC-----KESVSVTS-----PSGVSS-----	50	
		: : : : : :		
Db	31	STDRGHLHLVYELVDSYDEVLNGYCNELDIYTIINKGSIISIEDNGKMGPTGIHKSKPT	90	
QY	51	-----TSGVSTSTMNHS-----LLQEKHR---ILHKLLONGNS	82	
		: : : : : :		
Db	91	VEVIFTVLHAGKPGCGGKYKTCGLGHCVASVNALSEMLEVEIIRHGRNITVYHOSFKNGGS	150	
QY	83	PAL--VAKITQAATGKSDSTTSCGCGDNVYKQ-----EQLSPKKKKNNALLRL	129	
		: : : : : :		
Db	151	PSGGLVKKGKTKKKG--TKVTERPDPTIFASTSFNFVDISERLQASDLKLNK	202	

RESULT 13
 US-08-158-353-4
 : Sequence 4, Application US/08158353
 Patent No. 5620862
 : GENERAL INFORMATION:
 : APPLICANT: Padula, Steven J.
 : TITLE OF INVENTION: Methods for Diagnosing Early Lyme
 : TITLE OF INVENTION: Disease
 : NUMBER OF SEQUENCES: 7
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
 : STREET: Two Millitia Drive
 : CITY: Lexington
 : STATE: MA
 : COUNTRY: USA
 ZIP: 02173
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/158,353
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Carroll, Alice O.
 REGISTRATION NUMBER: 33,542
 REFERENCE/DOCKET NUMBER: UCT93-05
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9340
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 212 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-158-353-4

Query Match	10.7%;	Score 75;	DB 1;	Length 212;
Best Local Similarity	20.6%;	Pred. NO. 1.4;		
Matches 29;	Conservative 25;	Mismatches 49;	Indels 38;	Gaps 4;

```

0Y    9 LLOJLTTSSDDRGHSLILNPSLBDSCK-----ESSVSTSS-----45
      | : : : : |
Dd    13 LELEFICSNNSGKGSDASTNPADBSAAGPNLTETSKRITSSNAFLVAKEVEYLVLISIDE 72
      | : : : : |
0Y    46 -----CVSSSTSGGVASTSMHGSLIGKHRIT-----LHKELONGNSPAEYAKITQAQ 92
      | : : : : | : : : : | : : : : |
Db    73 LAKKAIGOKIDNNNGALALANQNSSLAGAIAISTLTTEKLKSLKLNBEULKETELAK--AK 130
      | : : : : | : : : : | : : : : |
0Y    93 ATGKDTSSITSCGDGNVVKOE 113
      | : : : : | : : : : |
Db    131 KCSEEPFTKLKLSGHADLGKOD 151

```

RESULT 14
US-09-046-

```

: Sequence 35, Application US/09046894
: Patent No. 6190857
:
: GENERAL INFORMATION:
:
: APPLICANT: Ralph, David
: APPLICANT: An, Gang
: APPLICANT: O'Hara, Mark S.
: APPLICANT: Veltiri, Robert
: TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA
: TITLE OF INVENTION: PROFILES IN PERIPHERAL LEUKOCYTES
:
: NUMBER OF SEQUENCES: 55
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77210
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/046,894
: FILING DATE: Concurrently Herewith
: CLASSIFICATION:
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,576
: FILING DATE: 24-MAR-1997
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Nakashima, Richard A.
: REGISTRATION NUMBER: P-42,023
: REFERENCE/DOCKET NUMBER: UROC:014
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (512) 418-3000
: TELEFAX: (512) 474-7577
:
: INFORMATION FOR SEQ ID NO: 35:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 761 amino acids
: TYPE: amino acid
: STRANDEDNESS:
:
: TOPOLOGY: linear
:
: US-09-046-894-35

```

Query Match 10.6% Score 74.5 DB 4 Length 761;
Best Local Similarity 27.3% Pred. No.9.9;
Matches 36; Conservative 17; Mismatches 42; Indels 37; Gaps 5

OY 14 TCSSSD-----KQHSLSLNSPLDSCKRESVSYTSPGVSSTFSGVYSTSNMHGSL 66
||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 599 TCCSSDALGPMSTEEHGLAISPERSRPSCHADFLITP--LSPTASQG--SSPCFGVSG 654

Qy	67	QEK-HRIHLKTLONGNSPAEYAKTTAQTGDTSSITSCGGNVNRQEQLSPKKKENAL	125
Db	655	EDPPEPPFAQGLRVNGKAKADVMPXTA-----PKKDERSLV	689
Qy	126	LRYLLDRDDPSD	137
Db	690	PAPAVDSGDSD	701

RESULT 15
US-08-933-750C-45

```

: Sequence 45, Application US/08933750C
: Patent No. 5932442
:
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Purvi
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,750C
: FILING DATE: September 23, 1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
:
: INFORMATION FOR SEQ. ID NO.: 45:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 811 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: LIVERUT04
: CLONE: 2515476
:
: US-08-933-750C-45

```

[illegible]

Db	705	EDSPFSEFAQMLRGKAKADVMPXTA-----		PKDENSELV 739
Qy	126	LRYLDRDDPSD 137		
Db	740	PPAPVDSGSED 751		

Search completed: April 19, 2001, 16:50:51
Job time: 299 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 16:54:00 ; Search time 147.68 Seconds
(without alignments)
152.489 Million cell updates/sec

Title: US-09-041-994-2_COPY_613_752

Perfect score: 703

Sequence: 1 LESKGHKKLLQLLTCSDDR.....ENNALLRYLLDRDDPSDALSL 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main.*
1: /cgn2_6/ptodata/2/paa/PCTRUS_COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pcp.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pcp.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pcp.*
23: /cgn2_6/ptodata/2/paa/US060_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	703	100.0	1415	14	US-09-041-994-2
2	703	100.0	1415	19	US-09-513-066-2
3	700	99.6	951	15	US-09-125-635-8
4	700	99.6	1420	15	US-09-125-635-8
5	628.5	89.4	1398	19	US-09-513-066-14
6	628.5	89.4	1402	15	US-09-125-635-12
7	628.5	89.4	1402	18	US-09-445-353A-2
8	540	76.8	149	23	US-60-236-359-19468
9	482	68.6	1391	19	US-09-513-066-15
10	292.5	41.6	1127	23	US-60-258-273-113

11	292.5	41.6	1464	12	US-08-891-640-2
12	287.5	40.9	151	23	US-60-182-918-63
13	286.5	40.8	1462	18	US-09-445-353A-3
14	192.5	27.4	193	23	US-60-177-667-86
15	192.5	27.4	193	23	US-60-177-667-87
16	188.5	26.8	48	12	US-08-891-640-5
17	186.5	26.5	1036	12	US-08-891-640-3
18	186.5	26.5	1061	11	US-08-701-154A-5
19	181	25.7	75	23	US-60-196-714-529
20	172	24.5	34	16	US-09-281-717-15
21	172	24.5	34	16	US-09-281-717-18
22	168	23.9	34	16	US-09-281-717-21
23	141	20.1	27	1	PCT-US99-24956-5
24	141	20.1	27	1	PCT-US99-24956-5
25	141	20.1	27	1	PCT-US99-30947-5
26	141	20.1	32	16	US-09-281-717-19
27	141	20.1	32	16	US-09-281-717-22
28	139	19.8	34	16	US-09-281-717-17
29	135	19.2	34	16	US-09-281-717-14
30	135	19.2	34	16	US-09-281-717-20
31	134	19.1	32	16	US-09-281-717-16
32	128	18.2	25	1	PCT-US99-24956-2
33	128	18.2	25	1	PCT-US99-24956-2
34	128	18.2	25	1	PCT-US99-30947-2
35	128	18.2	25	19	US-09-513-066-26
36	106	15.1	34	16	US-09-281-717-9
37	105	14.9	34	16	US-09-281-717-6
38	105	14.9	34	16	US-09-281-717-12
39	97	13.8	491	15	US-09-107-532-6115
40	97	13.8	1113	20	US-09-618-425-9
41	95	13.5	18	19	US-09-513-066-24
42	93.5	13.3	25	1	PCT-US99-24956-3
43	93.5	13.3	25	1	PCT-US99-24956-3
44	93.5	13.3	25	1	PCT-US99-30947-3
45	93.5	13.3	1357	23	US-60-173-464-29513

ALIGNMENTS

RESULT 1
US-09-041-994-2
; Sequence 2, Application US/09041994
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Li, Hui
; TITLE OF INVENTION: Transcriptional Coactivator for Nuclear
; TITLE OF INVENTION: Hormone Receptors
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive and Cockfield
; STREET: 28 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,994
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Liepmann, W. Hugo
; REGISTRATION NUMBER: 20,407
; REFERENCE/DOCKET NUMBER: UMM-026-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-742-4214
; INFORMATION FOR SEQ ID NO: 2:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-994-2

Query Match          100.0%; Score 703; DB 14; Length 1415;
Best Local Similarity 100.0%; Pred. No. 7.8e-60;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 613 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 672

QY 61 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 120
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 673 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 732

QY 121 ENNALLRYLLDRDDPSDALS 140
    ||||||||||||||||||
Db 733 ENNALLRYLLDRDDPSDALS 752

RESULT 2
US-09-513-066-2
; Sequence 2, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Leo, Christopher
; APPLICANT: Li, Hui
; TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
; FILE REFERENCE: UMG-046CP
; CURRENT APPLICATION NUMBER: US/09/513,066
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: USSN 09/041,994
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: USSN 60/073,674
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-066-2

Query Match          100.0%; Score 703; DB 19; Length 1415;
Best Local Similarity 100.0%; Pred. No. 7.8e-60;
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 613 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 672

QY 61 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 120
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 673 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 732

QY 121 ENNALLRYLLDRDDPSDALS 140
    ||||||||||||||||||
Db 733 ENNALLRYLLDRDDPSDALS 752

RESULT 3
US-09-125-635-8
; Sequence 8, Application US/09125635
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
```

```
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; PRIOR FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-125-635-8

Query Match          99.6%; Score 700; DB 15; Length 951;
Best Local Similarity 99.3%; Pred. No. 8.8e-60;
Matches 139; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 314 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 373

QY 61 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 120
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 374 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 433

QY 121 ENNALLRYLLDRDDPSDALS 140
    ||||||||||||||||||
Db 434 ENNALLRYLLDRDDPSDALS 453

RESULT 4
US-09-125-635-4
; Sequence 4, Application US/09125635
; GENERAL INFORMATION:
; APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; PRIOR FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-125-635-4

Query Match          99.6%; Score 700; DB 15; Length 1420;
Best Local Similarity 99.3%; Pred. No. 1.5e-59;
Matches 139; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 613 LESKGHKLLQLLTCSSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSTSN 672

QY 61 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 120
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 673 MHGSLLOEKHRIHLKLLQNGNSPAEVAKITAQATGKDTSSITSCGNGNVYKQQLSPKKK 732

QY 121 ENNALLRYLLDRDDPSDALS 140
    ||||||||||||||||||
Db 733 ENNALLRYLLDRDDPSDALS 752

RESULT 5
US-09-513-066-14
; Sequence 14, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
```

APPLICANT: Leo, Christopher
APPLICANT: Li, Hui
TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
STERIOD NUCLEAR RECEPTORS
FILE REFERENCE: UMG-026CP
CURRENT APPLICATION NUMBER: US/09/513,066
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: USSN 09/041,994
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: USSN 60/073,674
PRIOR FILING DATE: 1998-02-04
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 1398
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-066-14

Query Match 89.4%; Score 628.5; DB 19; Length 1398;
Best Local Similarity 89.3%; Pred. No. 1.6e-52;
Matches 125; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
QY 1 LESKGHKKLLQLLTCCSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSSTSN 60
DB 606 LESKGHKKLLQLLTCCSDDRGHSSLTNSPLDPCNCKDSSVSVTSPSGVSSSTSGVSSSTSN 665
QY 61 MHGSLLOEKHRLHLKLLONGNSPAEVAKITAGATGKDTSSITSCDGNVVKOEQLSPKKK 120
DB 666 VHGSLLQEKHRLHLKLLONGNSPAEVAKITAEATGKDTSSITSCGEG-TTROEQLSPKKK 724
QY 121 ENNALLRYLLDRDPDSALS 140
DB 725 ENNALLRYLLDRDPDSVLA 744

RESULT 6
US-09-125-635-12
Sequence 12, Application US/09125635
GENERAL INFORMATION:
APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE
TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
FILE REFERENCE: 49944
CURRENT APPLICATION NUMBER: US/09/125,635
CURRENT FILING DATE: 1998-08-21
PRIOR APPLICATION NUMBER: 60/049,728
PRIOR FILING DATE: 1997-06-17
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 1402
TYPE: PRT
ORGANISM: Mus musculus
US-09-125-635-12

Query Match 89.4%; Score 628.5; DB 15; Length 1402;
Best Local Similarity 89.3%; Pred. No. 1.6e-52;
Matches 125; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
QY 1 LESKGHKKLLQLLTCCSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSSTSN 60
DB 605 LESKGHKKLLQLLTCCSDDRGHSSLTNSPLDPCNCKDSSVSVTSPSGVSSSTSGVSSSTSN 664
QY 61 MHGSLLOEKHRLHLKLLONGNSPAEVAKITAGATGKDTSSITSCDGNVVKOEQLSPKKK 120
DB 665 VHGSLLQEKHRLHLKLLONGNSPAEVAKITAEATGKDTSSITSCGEG-TTROEQLSPKKK 723
QY 121 ENNALLRYLLDRDPDSALS 140
DB 724 ENNALLRYLLDRDPDSVLA 743

RESULT 7
US-09-445-353A-2
Sequence 2, Application US/09445353A
GENERAL INFORMATION:
APPLICANT: Rosenfeld, et al., Michael
TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
FILE REFERENCE: 6627-PAL021
CURRENT APPLICATION NUMBER: US/09/445,353A
CURRENT FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1402
TYPE: PRT
ORGANISM: Mus musculus
US-09-445-353A-2

Query Match 89.4%; Score 628.5; DB 18; Length 1402;
Best Local Similarity 89.3%; Pred. No. 1.6e-52;
Matches 125; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
QY 1 LESKGHKKLLQLLTCCSDDRGHSSLTNSPLDSSCKESSVSVTSPSGVSSSTSGVSSSTSN 60
DB 606 LESKGHKKLLQLLTCCSDDRGHSSLTNSPLDPCNCKDSSVSVTSPSGVSSSTSGVSSSTSN 665
QY 61 MHGSLLOEKHRLHLKLLONGNSPAEVAKITAGATGKDTSSITSCDGNVVKOEQLSPKKK 120
DB 666 VHGSLLQEKHRLHLKLLONGNSPAEVAKITAEATGKDTSSITSCGEG-TTROEQLSPKKK 724
QY 121 ENNALLRYLLDRDPDSALS 140
DB 725 ENNALLRYLLDRDPDSVLA 744

RESULT 8
US-60-236-359-19468
Sequence 19468, Application US/60236359
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
FILE REFERENCE: HdbMORF-4P
CURRENT APPLICATION NUMBER: US/60/236,359
CURRENT FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 21709
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 19468
LENGTH: 149
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL034418.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
OTHER INFORMATION: EST_HUMAN HIT: AW503706.1, EVALUE 9.00e-51
OTHER INFORMATION: SWISSPROT HIT: Q58434, EVALUE 5.30e-01
US-60-236-359-19468

Query Match 76.8%; Score 540; DB 23; Length 149;
Best Local Similarity 99.1%; Pred. No. 3.4e-45;
Matches 108; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 32 SSKKESVSVTSPSGVSSSTSGGVSSTNNHGSLLQEKHRLHKLKLLONGNSPAEVAKITA 91
|||||
Db 1 SSKKESVSVTSPSGVSSSTSGGVSSTNNHGSLLQEKHRLHKLKLLONGNSPAEVAKITA 60
|||||

QY 92 QATGKDTSSITSCGDGNVVKQEOQLSPKKENNALLRYLLDRDDPSDALS 140
:|||||
Db 61 EATGKDTSSITSCGDGNVVKQEOQLSPKKENNALLRYLLDRDDPSDALS 109
:|||||

RESULT 9
US-09-513-066-15
; Sequence 15, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Leo, Christopher
; APPLICANT: Li, Hui
; TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
; TITLE OF INVENTION: STEROID NUCLEAR RECEPTORS
; FILE REFERENCE: UMG-026CP
; CURRENT APPLICATION NUMBER: US/09/513,066
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: USSN 09/041,994
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: USSN 60/073,674
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-066-15

Query Match 68.6%; Score 482; DB 19; Length 1391;
Best Local Similarity 72.0%; Pred. No. 3.9e-38;
Matches 103; Conservative 17; Mismatches 15; Indels 8; Gaps 4;

QY 2 ESKGKKLLQLLTCCSDDRGHSSLTNSPLDSSCKESSVVTSPSGVSSSTSGGVSSTSNM 61
|||||
Db 612 ESKGKKLLQLLTCTEERGQSLMSSSDM--CKDSS-VNTSPSGVSSSTSGGVSSTSNL 668
|||||

QY 62 HGSLLQEKHRLHKLKLLONGNSPAEVAKITAQATGKD---TSSITSCGDGNVVKQEOQLSP 117
|||||

Db 669 HGSMLQEKHRLHKLKLLONGNSPAEVAKITAATGKDFQETVSSAPCTEA-TVKREQLSP 727
|||||

QY 118 KKENNALLRYLLDRDDPSDALS 140
|||||

Db 728 KKENNALLRYLLDRDDPSDALS 750
|||||

RESULT 10
US-60-258-273-113
; Sequence 113, Application US/60258273
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
; FILE REFERENCE: GL001042-PROV
; CURRENT APPLICATION NUMBER: US/60/258,273
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 1127
; TYPE: PRT
; ORGANISM: Human

US-60-258-273-113

Query Match 41.6%; Score 292.5; DB 23; Length 1127;
Best Local Similarity 48.2%; Pred. No. 1.2e-19;
Matches 67; Conservative 22; Mismatches 31; Indels 19; Gaps 4;

QY 2 ESKGKKLLQLLTCCSDDRGHSSLTNSPLDSSCKESSVVTSPSGVSSSTSGGVSSTSNM 61
:|||||
Db 354 DSKGQTKLLQLLTCTKSD-----QMEPSPLASSL-----SDTNKDTGSLPGSGST 398
|||||

QY 62 HGSLLQEKHRLHKLKLLONGNSPAEVAKITAQATGKD---TSSITSCGDGNVVKQEOQLSPK 118
|||||

Db 399 HGTSLKEKHRLHKLKLLQDSSSPVDLAKLTAEATGKDLQESSSTAPGSEVTIQEPPVSPK 458
|||||

QY 119 KKENNALLRYLLDRDDPSD 137
|||||

Db 459 KKE-NALLRYLLDKDPTKD 476
|||||

RESULT 11
US-08-891-640-2
; Sequence 2, Application US/08891640
; GENERAL INFORMATION:
; APPLICANT: Chambon, Pierre
; APPLICANT: Gronemeyer, Hinrich
; APPLICANT: Voegel, Johannes
; APPLICANT: Lutz, Yves
; TITLE OF INVENTION: Transcriptional Intermediary Factor-2
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,640
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/021,247
; FILING DATE: 12-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1383.0130001/EKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-891-640-2

Query Match 41.6%; Score 292.5; DB 12; Length 1464;
Best Local Similarity 48.2%; Pred. No. 1.8e-19;
Matches 67; Conservative 22; Mismatches 31; Indels 19; Gaps 4;

QY 2 ESKGKKLLQLLTCCSDDRGHSSLTNSPLDSSCKESSVVTSPSGVSSSTSGGVSSTSNM 61
:|||||
Db 634 DSKGQTKLLQLLTCTKSD-----QMEPSPLASSL-----SDTNKDTGSLPGSGST 678
|||||

Qy	Db	Qy	Db
62	HGSLHGEHRIHKLHNGNSPAEVAKITQATKCD--TSSIFSCGDGNNVKEOLSPK 118	119	KKENALLRTYLDNDPSPD 137
679	HGTSLEKHKRLHRLLDSSSPVDLAKLTAEATFOKDLSSSSSTAPSEVTINKEPVSFK 728	739	KKE--NALRTYLLDKDPTKD 756

```

RESULT 12
US-60-182-918-63
: Sequence 63, Application US/60182918
: GENERAL INFORMATION:
: APPLICANT: Bonazzi, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
: TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
: FILE REFERENCE: CL000266
: CURRENT APPLICATION NUMBER: US/60/182,918
: CURRENT FILING DATE: 2000-02-16
: NUMBER OF SEQ. ID NOS.: 64
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ. ID NO. 63
: LENGTH: 151
: TYPE: PRT
: ORGANISM: Human
: US-60-182-918-63

```

[illegible]

```

RESULT 13
US-09-445-353A-3
; Sequence 3, Application US/09445353A
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, et al., Michael
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/Clp
; FILE REFERENCE: 6627-P/1021
; CURRENT APPLICATION NUMBER: US/09/445,353A
; CURRENT FILING DATE: 1998-06-12
; NUMBER OF SEQ. ID NOS.: 3
; SOFTWARE: Patentin version 3.0
; SEQ. ID NO 3
; LENGTH: 1462
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-445-353A-3

```

	Query Match	Score	DB 18	Length	1462
	Best Local Similarity	48.2%	Pred. No. 6,8e-19		
Matches	67	Conservative	20	Mismatches	33
				Indels	19
				Gaps	4
OY	2	ESKGRKLLQLLTCTSSDRGHSLNPLDSSCKESSVYTPSGVSSSTSGCVSSTSMK	61		
Db	634	DSKGGTKLLQLLTCTSSDRGHSLNPLDSSCKESSVYTPSGVSSSTSGCVSSTSMK	678		
OY	62	HGSLLEQKHRIHLKLLQNGNSPFAEVAKITVAQTGKD--TSSITSCGCGDNVYKQEOELSPK	118		

Db 679 HGTSKEKHKILHRLDSSSPVLAKTATKELSQSSSSTAPSEVTVAKQEPASPK 738

Qy 119 KKNENALRYLLDRDDPSD 137

Db 739 KKE-NALLRYLLDDKDDTFD 756

```

RESULT 14
US-60-177-667-86
: Sequence 86, Application US/60177667
: GENERAL INFORMATION:
: APPLICANT: Bonazzi, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
: TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
: FILE REFERENCE: CL000197
: CURRENT APPLICATION NUMBER: US/60-177,667
: CURRENT FILING DATE: 2000-01-27
: NUMBER OF SEQ ID NOS: 140
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 86
: LENGTH: 193
: TYPE: PRT
: ORGANISM: HUMAN
: US-60-177-667-86

```

	Query Match Similarity	27.4%; Score 192.5; DB 23;	Length 193;
Best Local Similarity	38.1%;	Pred No.6.8e-11;	
Matches	53; Conservative	33; Mismatches	36; Indels 15; Gaps 7.
Oy	3 SKGAKRLQLLQTLCCSSDDRRGHSLITMSPDLSCKEKESSVYVTSPTSGGVSSSTGGVVSTRNMH : : : : : : : : : : : : : : : : : : : Db 16 SGTSHKVLQGLLTTFTEAQ---LRHADIDTSCKD-VLSCTGTGNASMSAGSQCSSL-H 69		
Oy	63 GSLLDEKRIRILKLIONGNSPAEVAVKIYAQAATSKDKDTSTSITSCGDENV-----VKOQLSP 117 : : : : : : : : : : : : : : : : : : : Db 70 SS-LTERKHIIHLRLOEG-SPSDITTLISVEPKKDSASTSVSVTGVOQNSSIKLDELAS 127		
Oy	118 KKKE--NNALLPYLLDRDD 134 :: : Db 128 KRESKDHQLLRYLLDKDE 146		

```

RESULT 15
US-60-177-667-87
: Sequence 87, Application US/60177667
:
: GENERAL INFORMATION:
: APPLICANT: Bonazzel, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
: FILE REFERENCE: CL000197
: CURRENT APPLICATION NUMBER: US/60-177, 667
: CURRENT FILING DATE: 2000-01-27
: NUMBER OF SEQ ID NOS: 140
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 87
:
: LENGTH: 193
:
: TYPE: PRT
:
: ORGANISM: HUMAN
US-60-177-667-87

```

Query Match 27.4%; Score 192.5; DB 23; Length 193;
 Best Local Similarity 38.1%; Pred. No. 6, 8e-11;
 Matches 53; Conservative 33; Mismatches 39; Indels 15; Gaps 7

OY 3 SKGAKKLLQLTCSDDRGHSLLNSPLDSCKRESKESVYTFSPGYSSTSGCVSSTSNMH 62
 16 SQTSHKLVQLLTYYYEQQ---LRAADIDITDCKD-VLSTGTGNSAKSNAVSSGSCPS--H 69

QY 63 GSLQEKHRIILHKILQNGNSPAEVAKITAQATGKDTSSITSCGDNV-----VKQQLSP 117
Db 70 SS-LTERHKILHRLQEG-SPSDITTTLSVEPDKKDSASTSVVTGQYQGNSSIKLELDAS 127
QY 118 KKE--NNALLRYLLDRDD 134
Db 128 KKEKSDHQLRYLLDKDE 146

Search completed: April 19, 2001, 16:54:03
Job time: 490 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 16:54:07 ; Search time 7.55 Seconds
(without alignments)
46.942 Million cell updates/sec

Title: US-09-041-994-2_COPY_613_752

Perfect score: 703
Sequence: 1 LESKGHKKLLQLLTCCSDR.....ENNALLRYLLDRDDPSDALS 140

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 12259 seqs, 2531507 residues

Total number of hits satisfying chosen parameters: 12259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PC7_NEW.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US60_NEW.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	11.1	1589	5	US-09-421-124-189
2	78	11.1	1589	5	US-09-421-124-189
3	70.5	10.0	826	5	US-09-533-077-330
4	70	10.0	509	5	US-09-739-449-11239
5	66.5	9.5	296	5	US-09-739-449-9535
6	65	9.2	382	5	US-09-739-449-8471
7	65	9.2	787	5	US-09-739-449-9028
8	63.5	9.0	2000	6	US-60-248-505-714
9	63	9.0	775	5	US-09-739-449-10931
10	63	9.0	1833	4	US-08-945-567C-4
11	63	9.0	1992	4	US-08-945-567C-3
12	63	9.0	2329	5	US-09-421-124-16
13	63	9.0	2329	5	US-09-421-124-16
14	63	9.0	3418	5	US-09-421-124-44
15	63	9.0	3418	5	US-09-421-124-44
16	62.5	8.9	353	5	US-09-293-858B-23
17	62.5	8.9	539	5	US-09-533-077-376
18	62	8.8	751	5	US-09-739-449-9898
19	61	8.7	373	5	US-09-819-142-2
20	61	8.7	316	5	US-09-739-449-13067
21	61	8.7	902	6	US-60-248-505-973
22	61	8.7	992	6	US-60-248-505-1112
23	60.5	8.6	580	6	US-60-248-505-1300
24	60.5	8.6	661	6	US-60-248-505-739
25	60.5	8.6	1731	6	US-09-739-449-8331
26	59.5	8.5	152	1	PCT-US01-01350-146
27	59.5	8.5	213	5	US-09-739-449-11798

28	59.5	8.5	315	6	US-60-277-380-40	Sequence 40, App1
29	59.5	8.5	414	5	US-09-739-449-11691	Sequence 11691, A
30	59.5	8.5	934	6	US-60-248-505-1027	Sequence 1027, Ap
31	59.5	8.5	933	5	US-09-207-359B-8	Sequence 8, App1
32	59	8.4	198	5	US-09-739-449-10949	Sequence 10949, A
33	59	8.4	537	6	US-60-248-505-1227	Sequence 1227, Ap
34	59	8.4	4563	5	US-09-802-640-32	Sequence 32, App1
35	58.5	8.3	269	5	US-09-792-024-114	Sequence 114, App
36	58.5	8.3	284	6	US-60-248-505-906	Sequence 906, App
37	58.5	8.3	352	5	US-09-514-450-5	Sequence 5, App1
38	58.5	8.3	587	6	US-60-248-505-742	Sequence 742, App
39	58.5	8.3	646	6	US-60-248-505-935	Sequence 935, App
40	58	8.3	310	5	US-09-739-449-10889	Sequence 10889, A
41	58	8.3	343	5	US-09-739-449-9832	Sequence 9832, Ap
42	57.5	8.2	608	5	US-09-739-449-10801	Sequence 10801, A
43	57.5	8.2	715	5	US-09-792-024-85	Sequence 85, App1
44	57	8.1	274	5	US-09-813-408-17	Sequence 17, App1
45	57	8.1	508	5	US-09-563-997-4	Sequence 4, App1

ALIGNMENTS

RESULT 1
US-09-421-124-189
Sequence 189, Application US/09421124
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Mooster, Richard F
APPLICANT: Ashworth, Alan
APPLICANT: Stratton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESS: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO.: 189:
SEQUENCE CHARACTERISTICS:
LENGTH: 1589 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-421-124-189
Query Match 11.1%; Score 78; DB 5; Length 1589;

[illegible]

```

Oy      60 NMHGSLLQEKHRILH-----LONGNSPAEAKI-----89
Db      1379 ---GALTVVESALKKVACMLEEPDLINTHTLQSPILPEVSKILPPQCAIRTPYV 1435
Oy      90 -TAQVTGKDTSSI-----TSCGDGNVVKOE--QLSPKKENMALL 126
Db      1436 NSLQKTYNDKSSLPSENVKESGSSGNTQSLQVSLQSLQSNRNDQTL 1481

RESULT      3
US-09-533-077-330
; Sequence 330, Application US/09533077
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C5
; CURRENT APPLICATION NUMBER: US/09/533.077
; CURRENT FILING DATE: 2000-03-22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 826
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-533-077-330

Query Match      10.0%; Score 70.5; DB 5; Length 826;
Best Local Similarity 26.4%; Pred. No. 4.6;
Matches 28; Conservative 19; Mismatches 44; Indels 15; Gaps 4;

Oy      28 SPLDSS-----CKESSVYTPSGVSSSTSGVSSFTSNMHGSLLOEKHRIH-KLLQN 79
Db      581 SPLESSASPEASAPQSVTVYEQIQIQIEPTANATTATTATDELKTIVKDMEDIKILIA 640
Oy      80 GNSPKEVAKITRAQTG---KDTSSITS---CGGDSNVYKQEOLSK 118
Db      641 SPSPPHIKETTSATSSPYRDTQSRTPASPNRAGKGVIEQTERKSHPR 686

RESULT      4
US-09-739-449-11239
; Sequence 11239, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-101154901C
; CURRENT APPLICATION NUMBER: US/09/739.449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 11239
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-11239

Query Match      10.0%; Score 70; DB 5; Length 509;
Best Local Similarity 25.9%; Pred. No. 2.8;
Matches 36; Conservative 20; Mismatches 59; Indels 24; Gaps 6;

Oy      6 HKKLLQLTCCSDDDKGHSLTNSPLDSSCKESSVSVTPSGVSSSTSG-----G 54
Db      302 HSRLLERAKUSDEMGAGSLTALPVIEI-QGNDVSALFIPYVISTDQIFLETFDLYQG 360

```

OY 55 VSTSNMHSLLQEKRIHLKILON--GNSPAEVAKITAQAT---GKDTSSITS--CGD 106
DB 361 IRPAVAVGSLSVRSVSAQAQKAMKQVAGSIKGEIAQREMAFAQGSIDLASTORLNR 420
OY 107 G----NVVKOEQLSPKKE 121
DB 421 GARTELELKQPOESPLKTE 439

RESULT 5

US-09-739-449-9535
; Sequence 9535, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9535
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9535

Query Match 9.5%; Score 66.5; DB 5; Length 296;
Best Local Similarity 23.9%; Pred. No. 3.2;
Matches 28; Conservative 24; Mismatches 46; Indels 19; Gaps 6;

OY 4 KHKKLLQLTSSDDRGHSSL-TNSPLDSCCKSSSVTSBGSST---SGGVSSST 58
DB 26 KGVKPLAVVIVGNDPASHAYVNSKMAKQCGFNSIQHTLPEETQALMLKLVGELNTD 85
OY 59 SNMHGSLQ---EKH---RILHKLQ-----NNSPAEVAKITAQGTGDTSSITSC 104
DB 86 ASIHGILVOLPLPKRHNSEDTIIOISILPEKVDGLSVLNAKGL---ATGDLATGLLISC 139

RESULT 6

US-09-739-449-8471
; Sequence 8471, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 8471
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-8471

Query Match 9.2%; Score 65; DB 5; Length 382;
Best Local Similarity 23.1%; Pred. No. 6.1;
Matches 33; Conservative 23; Mismatches 53; Indels 34; Gaps 9;

OY 16 SSDDRGHSSLTNSPLDSCCKSSSVTSBGSST---VSSSTGCGVSTSNMHGSLQEK-HRI 72
DB 2 TTDQTPHGFDPDOPPLASEMERLILDIYKRHRPIARSSITG-----HTNLQGSVHRU 53
OY 73 LKHLQNG-----NSPAEVAKITAQAT---GKDTSSITSC-GD--GNVVKO 112
DB 54 IEALDGRGLQGTGAPLKTGRGQSPTEIEVPEAAVSLGISINTDSVVICIADPRCAVIVE 113

OY 113 EQLS--PKKKEN-ALLRYLDR 132
DB 114 EKLKMLPEDERETTALALSOALDR 136

RESULT 7

US-09-739-449-9028
; Sequence 9028, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9028
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9028

Query Match 9.2%; Score 65; DB 5; Length 787;
Best Local Similarity 35.0%; Pred. No. 15;
Matches 21; Conservative 7; Mismatches 28; Indels 4; Gaps 1;

OY 14 TCSDDRGHSSLTNSPLDSCCKSSSVTSBGSSTSGVSSSTSNMHGSLQEKHRL 73
DB 419 TYSSSPSTSSASASPSGSSMTSTSPATTTGCASTPTASVSSP---GSAKKGHSIL 474

RESULT 8

US-60-248-505-714
; Sequence 714, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 714
; LENGTH: 2000
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-714

Query Match 9.0%; Score 63.5; DB 6; Length 2000;
Best Local Similarity 23.6%; Pred. No. 63;
Matches 39; Conservative 24; Mismatches 51; Indels 51; Gaps 7;

OY 3 SKGKRLQLLT-----CSSDDRGHSSLTNSPLDSCCKSSSV----- 41
DB 159 SKYHREIAGLIGSEWTRIVYLSNRSSEASMPPTTSASVSLVLAQEGGSAASY 218
OY 42 TS-----PSGVSSSTGCGVSTSNMHGSLQEK-----KHTLHKLQ--NGN 81
DB 219 TSATBGRGPASGLATREDSGNGSSQNGSGSIRHEDQCDLRFRFLFYDLQELDQSLSGE 278
OY 82 SPAEVAKITAQATGKDTSSITSCGCGVVKQEDLSKKKENNAL 126
DB 279 APQ-----TQQAQEMLNINIESPGPSH-----QPTPHSENNSNL 313

RESULT 9


```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-421-124-16
```

Query Match 9.0%; Score 63; DB 5; Length 2329;
Best Local Similarity 18.6%; Pred. No. 85;

Matches 30; Conservative 29; Mismatches 56; Indels 46; Gaps 6;

```
Oy 3 SKGHKKLLQLTCCSDDRRHSSLTNSPLDSCK---ESSVSTSPSGVSSSTSGVSSSTS 59
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 982 SKCHDSVSMFKIEN---HNKTYSEKNNKCQLLQNNIEMTGTGFEIETENKRNTE 1037
Oy 60 N-----MHGS-----LLOEKHRLHLK-----LQNGNS 82
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1038 NEDNKTYAASRSHNLFEFGSDSSKNDYVCIRKDETDLLFTQHNICLKLSQGFMEKGMT 1097
Oy 83 PAEVAKITAQAATGKDTSSITSCGDGNVYKQEQLSPRKKENN 123
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1098 --QIKEDLSDLTFLEVAKAQEAHGNTSNKEQLTATKTEON 1136
```

RESULT 13

```
US-09-421-124-16
Sequence 16, Application US/09421124
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Straton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: NC 27622
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
APPLICATION NUMBER: GB 9523959.6
FILING DATE: 23-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9525555.0
FILING DATE: 14-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9617961.9
FILING DATE: 28-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kenneth D Sibley
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405-135
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2329 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-421-124-16
```

Query Match 9.0%; Score 63; DB 5; Length 2329;
Best Local Similarity 18.6%; Pred. No. 85;
Matches 30; Conservative 29; Mismatches 56; Indels 46; Gaps 6;

```
Oy 3 SKGHKKLLQLTCCSDDRRHSSLTNSPLDSCK---ESSVSTSPSGVSSSTSGVSSSTS 59
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 982 SKCHDSVSMFKIEN---HNKTYSEKNNKCQLLQNNIEMTGTGFEIETENKRNTE 1037
Oy 60 N-----MHGS-----LLOEKHRLHLK-----LQNGNS 82
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1038 NEDNKTYAASRSHNLFEFGSDSSKNDYVCIRKDETDLLFTQHNICLKLSQGFMEKGMT 1097
Oy 83 PAEVAKITAQAATGKDTSSITSCGDGNVYKQEQLSPRKKENN 123
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1098 --QIKEDLSDLTFLEVAKAQEAHGNTSNKEQLTATKTEON 1136
```

RESULT 14

```
US-09-421-124-44
Sequence 44, Application US/09421124
GENERAL INFORMATION:
APPLICANT: Futreal, Phillip A
APPLICANT: Wooster, Richard F
APPLICANT: Straton, Michael R
TITLE OF INVENTION: Materials and methods relating to the
TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer
NUMBER OF SEQUENCES: 222
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell Seltzer Park & Gibson
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107
CITY: Raleigh
STATE: NC
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,124
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/755,587
FILING DATE: 25-NOV-1996
```


This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 16:50:51 ; Search time 32.49 Seconds

(without alignments)
95.788 Million cell updates/sec

Title: US-09-041-994-2_COPY_1018_1179

Perfect score: 845
Sequence: 1 EQVSHGQNRLRLNSLDDL.....MRPRTPKQLRMQLQRLQ 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/2/1aa/PCRTUS_COMB.pep:*
- 6: /cgn2_6/prodata/2/1aa/ackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	10.5	415	3	US-09-176-657-8 Sequence 8, Appl1
2	85.5	10.1	3969	4	US-08-061-376-5 Sequence 5, Appl1
3	85	10.1	2414	1	US-08-227-536-2 Sequence 2, Appl1
4	85	10.1	2414	5	PCR-US95-04682-2 Sequence 2, Appl1
5	84	9.9	357	1	US-08-612-986-5 Sequence 5, Appl1
6	84	9.9	357	1	US-08-361-806A-5 Sequence 5, Appl1
7	84	9.9	357	5	PCR-US95-16806A-5 Sequence 5, Appl1
8	79	9.3	314	2	US-08-525-742-6 Sequence 6, Appl1
9	75.5	8.9	485	2	US-08-749-391-2 Sequence 2, Appl1
10	75.5	8.9	485	4	US-09-390-200-2 Sequence 2, Appl1
11	75	8.9	723	2	US-08-548-159-5 Sequence 5, Appl1
12	75	8.9	1078	3	US-08-480-474-11 Sequence 11, Appl1
13	74.5	8.8	510	1	US-08-278-635B-4 Sequence 4, Appl1
14	74.5	8.8	510	3	US-08-471-961-4 Sequence 4, Appl1
15	74.5	8.8	1863	2	US-08-603-753D-2 Sequence 2, Appl1
16	74.5	8.8	1863	4	US-09-099-753-2 Sequence 2, Appl1
17	74.5	8.8	1863	4	US-08-986-106-2 Sequence 2, Appl1
18	74.5	8.8	2441	1	US-08-194-468-2 Sequence 2, Appl1
19	74.5	8.8	2441	3	US-08-961-739-2 Sequence 2, Appl1
20	74	8.8	511	3	US-08-464-258B-4 Sequence 4, Appl1
21	73.5	8.7	241	2	US-08-867-087B-11 Sequence 11, Appl1
22	72.5	8.6	301	2	US-08-785-464-1 Sequence 1, Appl1
23	72.5	8.6	381	4	US-09-135-021-80 Sequence 80, Appl1
24	72.5	8.6	676	4	US-09-135-021-2 Sequence 2, Appl1
25	72.5	8.6	765	4	US-08-425-061-19 Sequence 19, Appl1
26	72.5	8.6	765	2	US-08-825-886-19 Sequence 19, Appl1
27	72.5	8.6	864	2	US-08-620-694A-2 Sequence 2, Appl1

28	72.5	8.6	864	3	US-09-022-255-2	Sequence 2, Appl1
29	72.5	8.6	864	3	US-09-022-696-2	Sequence 2, Appl1
30	72.5	8.6	864	3	US-08-978-773-2	Sequence 2, Appl1
31	72.5	8.6	864	3	US-09-022-253-2	Sequence 2, Appl1
32	72.5	8.6	864	3	US-09-022-260-2	Sequence 2, Appl1
33	72.5	8.6	864	4	US-09-022-259-2	Sequence 2, Appl1
34	72.5	8.6	864	4	US-09-022-257-2	Sequence 2, Appl1
35	72.5	8.6	900	1	US-08-425-061-20	Sequence 20, Appl1
36	72.5	8.6	900	2	US-08-825-886-20	Sequence 20, Appl1
37	72.5	8.6	914	1	US-08-425-061-21	Sequence 21, Appl1
38	72.5	8.6	914	2	US-08-825-886-21	Sequence 21, Appl1
39	72.5	8.6	1202	1	US-08-425-061-22	Sequence 22, Appl1
40	72.5	8.6	1202	2	US-08-825-886-22	Sequence 22, Appl1
41	72.5	8.6	1363	1	US-08-425-061-23	Sequence 23, Appl1
42	72.5	8.6	1363	2	US-08-825-886-23	Sequence 23, Appl1
43	72.5	8.6	1852	1	US-08-425-061-24	Sequence 24, Appl1
44	72.5	8.6	1852	2	US-08-825-886-24	Sequence 24, Appl1
45	72.5	8.6	1863	1	US-08-425-061-16	Sequence 16, Appl1

ALIGNMENTS

```
RESULT 1
US-09-176-657-8
; Sequence 8, Application US/09176657
; Patent No. 6020164
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Lu, Aina
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN RNA BINDING PROTEINS
; FILE REFERENCE: PF-0611 US
; CURRENT APPLICATION NUMBER: US/09/176,657
; CURRENT FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: g608464
US-09-176-657-8

Query Match 10.5%; Score 89; DB 3; Length 415;
Best Local Similarity 30.3%; Pred. No. 0.067;
Matches 40; Conservative 17; Mismatches 41; Indels 34; Gaps 7;

QY 36 LLDOLHTLSTNTDAGLEEDRALGI-----PELVNQGALRP--KODPFGOGAEAVMM 87
Db 242 ILDOQNPMLMOASLAIAQOQNNRALLRSPVLAOSLGRGPAIPGQOPAFQDAAALA-- 299
QY 88 DOKAGLYGQTPAGPQMGCFHLOGSPFSFNSMMNOMN--OQGNFPLOGMHRIMRP 145
Db 300 -----GNFLA---OLQGPPLFNAALQTNALQOSAF---GMDPNAVIALAL 339
QY 146 RYTPKQLRMQL 157
Db 340 LAN--EQARFQL 349

RESULT 2
US-08-061-376-5
; Sequence 5, Application US/08061376
; Patent No. 6175000
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Djabali, Malek
```

```

1  APPLICANT: Selter, Lucia
2  APPLICANT: Parry, Pauline
3  TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11Q23
4  TITLE OF INVENTION: TRANSLOCATION BREAKPOINT ASSOCIATED WITH ACUTE LEUKEMIAS
5  NUMBER OF SEQUENCES: 12
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
8  STREET: 444 South Flower Street, Suite 2000
9  CITY: Los Angeles
10 STATE: California
11 COUNTRY: USA
12 ZIP: 90071
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: Patentin Release #1.0, Version #1.25
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/061,376
21 FILING DATE: 13-MAY-1993
22 CLASSIFICATION: 435
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Reiter, Stephen E.
25 REGISTRATION NUMBER: 31,192
26 REFERENCE/DOCKET NUMBER: P41 9387
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (619)546-4737
29 TELEFAX: (619)546-9392
30 INFORMATION FOR SEQ ID NO: 5:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 3969 amino acids
33 TYPE: amino acid
34 STRANDEDNESS: unknown
35 TOPOLOGY: unknown
36 MOLECULE TYPE: protein
37
38 US-08-061-376-5
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: FDCI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-536-2

Query Match          10.1%; Score 85; DB 1; Length 2414;
Best Local Similarity 23.5%; Pred. No. 2.4;
Matches   44; Conservative 16; Mismatches   57; Indels   70; Gaps       7;

QY      6 GTGNPPLRLNSLDDLVGPSNLSEGSDEKALLDQLHTLLSTNDATNGLEIRALRIPELY 65
        ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB      2044 GVTSQAALQNLRLTRSPSSPLQQOQ---VLSIH-----ANPLL 2081

QY      66 NQGALPEEKDAFGQGEAEAVMDOKAGLYGOTYPACGP-----PMGG 108
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      2082 -----AATIKRAAYKANSNPGPIPGQPCMGPCGCGICPTMPGGQG 2123

QY      109 FHLCQSSEFSNMY-----MNMQNGNFPLQGMHPRANMRPTNT--PKOLR--M 155
        :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      2124 VHSNAMQNMMNQGVORAGLPQOQPQQQLPPMGGMSPPAQOMNMNHNTWPSQFRDL 2183

QY      156 QLOQRLO 162
        ::||::|
DB      2184 RKQMMQ 2190

RESULT      4
PCT-US95-04682-2
Sequence 2, Application PC/TUS9504682
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
```

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Helne, Ph. D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCT-308Xg999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

Query Match 10.1%; Score 85; DB 5; Length 2414;
Best Local Similarity 23.5%; Pred. No. 2.4;
Matches 44; Conservative 16; Mismatches 57; Indels 70; Gaps 7;

QY 6 GTGNRLRLSLDDLVGSPNLEGSDEKALLDQLTLSTNATGTEIDRALGPELV 65
DB 2044 GTVSQALQMLRLTLSPSSPLQOOO-VLSILH-----ANPOL 2081
QY 66 NCGALEPPKQDAFGQEAAMVMDOKAGLYGQYPAQGP-----PMGG 108
DB 2082 -----AAFIKORAKYANSNPDPTEGPGPMQPGQIGPPTMGQGG 2123
QY 109 FHLOGSPSPFSNM-----MNMNQGNFPLQGMHPRANIMRPTNT-PKOLR-M 155
DB 2124 VHSNPMQNNMPPMAGVQRAGLPGQDPQQLDPFMGMSPPQAQOMNNHNTMNSQFDDIL 2183
QY 156 QLOQRLLQ 162
DB 2184 RROQMMD 2190

RESULT 5
US-08-612-986-5
Sequence 5, Application US/08612986
Patent No. 5770384
GENERAL INFORMATION:
APPLICANT: Elliot J. Androphy
TITLE OF INVENTION: E2 BINDING PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,986
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/361,806
FILING DATE: 22 DEC 1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: NEP-004DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 357 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-612-986-5

Query Match 9.9%; Score 84; DB 1; Length 357;
Best Local Similarity 27.0%; Pred. No. 0.2;
Matches 33; Conservative 6; Mismatches 47; Indels 36; Gaps 3;

QY 73 PKDADFQGE-AAVMDOKAGLYGQYPAQGP----- 103
DB 199 PEHGQFGSGAYGTAPRPHYGPOTAPSPSQQLRAPSAPRAVQYLQOPORPAVHG 258
QY 104 ---PMGGFHLQGSFSPFSNMNMNQGNF-----PLOGMHPRANIMRPTNTPKOLRMD 156
DB 259 HFGPTQFGFQPGALSLQKOMEHANQGTGFSOSSSLRPMHQAALPRALLASPOLPQV 318
QY 157 LQ 158
DB 319 MQ 320

RESULT 6
US-08-361-806A-5
Sequence 5, Application US/08361806A
Patent No. 5792833
GENERAL INFORMATION:
APPLICANT: Elliot J. Androphy
TITLE OF INVENTION: E2 BINDING PROTEINS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/361,806A
FILING DATE: 22 DEC 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: NEP-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 357 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-361-806A-5

Query Match 9.9%; Score 84; DB 1; Length 357;
Best Local Similarity 27.0%; Pred. No. 0.2;
Matches 33; Conservative 6; Mismatches 47; Indels 36; Gaps 3;

QY 73 PKDADFQGE-AAVMDOKAGLYGQYPAQGP----- 103

TITLE OF INVENTION: Anaerobic Fungus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,391
FILING DATE: 13-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna M. Ferber
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-749-391-2

Query Match 8.9%; Score 75.5; DB 2; Length 485;
Best Local Similarity 37.9%; Pred. No. 2.7;
Matches 25; Conservative 2; Mismatches 32; Indels 7; Gaps 2;

QY 88 DQKAGLYGQTYPAQGPMPGCGFHLQGSFSPNSMMNQM-QQGNFPLQGHPRANITMRPR 146
DB 268 DQPGGPGGQPPGQGPQGGQPPGQGPQGGQPPGQNDQGGQPPGQGP-----PQ 321
QY 147 TNPQK 152
DB 322 GNDQOQ 327

RESULT 10
US-09-390-200-2
Sequence 2, Application US/09390200
Patent No. 6137032
GENERAL INFORMATION:
APPLICANT: Cheng, Kuo-Joan
APPLICANT: Selienger, Leonard B.
APPLICANT: Liu, Jin-Hao
APPLICANT: Hu, Youji
APPLICANT: Forsberg, Cecil W.
TITLE OF INVENTION: A xylanase Obtained From an
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/390,200
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,391
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Donna M. Ferber
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-390-200-2

Query Match 8.9%; Score 75.5; DB 4; Length 485;
Best Local Similarity 37.9%; Pred. No. 2.7;
Matches 25; Conservative 2; Mismatches 32; Indels 7; Gaps 2;

QY 88 DQKAGLYGQTYPAQGPMPGCGFHLQGSFSPNSMMNQM-QQGNFPLQGHPRANITMRPR 146
DB 268 DQPGGPGGQPPGQGPQGGQPPGQGPQGGQPPGQNDQGGQPPGQGP-----PQ 321
QY 147 TNPQK 152
DB 322 GNDQOQ 327

RESULT 11
US-08-548-159-5
Sequence 5, Application US/08548159
Patent No. 5989511
GENERAL INFORMATION:
APPLICANT: MacLaren, No. 5989511 K.
APPLICANT: No. 5989511kins, Abner L.
APPLICANT: Lau, Michael S.
APPLICANT: Li, Qing
TITLE OF INVENTION: Materials and Methods for Detection and
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/548,159
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UP154.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-375-5800
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:

DB 428 -----MEPETHPSQASEILLSPQIOKALE 452

RESULT 14
US-08-471-961-4

; Sequence 4, Application US/08471961

; Patent No. 6100046

; GENERAL INFORMATION:

; APPLICANT: ELGOYHEN, ANA BELEN

; APPLICANT: JOHNSON, DAVID S.

; APPLICANT: BOUTLER, JAMES R.

; APPLICANT: HEINEMANN, STEPHEN F.

; TITLE OF INVENTION: CLONING AND EXPRESSION OF A NOVEL

; TITLE OF INVENTION: ACETYLCHOLINE-GATED ION CHANNEL RECEPTOR

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GRAY CARY WARE & FREIDENRICH

; STREET: 4365 EXECUTIVE DRIVE, SUITE 1600

; CITY: SAN DIEGO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 92121

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/471,961

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/278,635

; FILING DATE: 21-JUL-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: REITER, STEPHEN E.

; REGISTRATION NUMBER: 31,192

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-677-1409

; TELEFAX: 619-677-1465

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 510 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-471-961-4

Query Match 8.8%; Score 74.5; DB 3; Length 510;

Best Local Similarity 22.7%; Pred. No. 3.7;

Matches 37; Conservative 13; Mismatches 36; Indels 77; Gaps 8;

QY 9 NRPILNLSDDLVPSPNLESGOSDERALLDQLHTLLSNLDATGLEGIDALGIPELVNG 68

DB 358 NRPL-----PPMELHGSPTDK-LSPSYHMETNMDA-----G 388

QY 69 QALEPKQDAFOGCEAV-----MMDOQKAG-LYGQTYPAQGPPOGPHLOGQSPFSNMNQ 124

DB 389 EREFTEEEBEDENCVCAGLPDSSMGVLYG-----HGLHLRA----- 427

QY 125 MNOQGNFPLQGHPRANIMRPTNTPKQ-----LRMOLOQRLO 162

DB 428 -----MEPETHPSQASEILLSPQIOKALE 452

RESULT 15
US-08-603-753D-2

; Sequence 2, Application US/08603753D

; Patent No. 5891857

; GENERAL INFORMATION:

APPLICANT: HOLT, JEFFREY T.

APPLICANT: JENSEN, ROY A.

APPLICANT: PAGE, DAVID L.

APPLICANT: KING, MARY-CLAIRE

APPLICANT: SZABO, CSILLA I.

APPLICANT: JETTON, THOMAS L.

APPLICANT: ROBINSON-BENTON, CHERYL L.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

APPLICANT: THOMPSON, MARILYN E.

```

?      AUTHORS:   Miki, Y., et al
?      TITLE:    A strong candidate gene for the breast and
?               ovarian cancer susceptibility gene BRCA1.
?      JOURNAL:   Science
?      VOLUME:   266
?      PAGES:    66-71
?      DATE:     1994
?      RELEVANT RESIDUES IN SEQ ID NO: 2: granin box
?      RELEVANT RESIDUES IN SEQ ID NO: domain at amino acids 1214-1223
; OS=08-603-f53d-2
```

[illegible]

Search completed: April 19, 2001, 16:50:54
Job time: 302 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2001, 16:54:03 ; Search time 147.68 Seconds

(without alignments)
176.452 Million cell updates/sec

Title: US-09-041-994-2_COPY_1018_1179

Perfect score: 845

Sequence: 1 EGVSHGTQNRPLRLNSLDDL.....MRPRTNPKQLRMQLQRLQ 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	845	100.0	951	US-09-125-635-8	Sequence 8, Appli
2	845	100.0	1415	US-09-041-994-2	Sequence 2, Appli
3	845	100.0	1415	US-09-513-066-2	Sequence 2, Appli
4	845	100.0	1420	US-09-125-635-4	Sequence 4, Appli
5	761	90.1	1402	US-09-125-635-12	Sequence 12, Appli
6	757	89.6	1398	US-09-513-066-14	Sequence 14, Appli
7	757	89.6	1402	US-09-445-353A-2	Sequence 2, Appli
8	564.5	66.8	1391	US-09-513-066-15	Sequence 15, Appli
9	312.5	37.0	1127	US-60-258-273-113	Sequence 113, App
10	312.5	37.0	1464	US-08-891-640-2	Sequence 2, Appli

11	310.5	36.7	1462	US-09-445-353A-3	Sequence 3, Appli
12	278	32.9	58	US-08-891-640-9	Sequence 9, Appli
13	266.5	31.5	1036	US-08-891-640-3	Sequence 3, Appli
14	266.5	31.5	1061	US-08-701-154A-5	Sequence 5, Appli
15	237	28.0	51	US-60-160-189-8436	Sequence 8436, Ap
16	237	28.0	51	US-60-169-867-5713	Sequence 5713, Ap
17	104	12.3	5222	US-60-245-201-208	Sequence 208, App
18	98.5	11.7	2247	US-60-167-217-16010	Sequence 16010, A
19	98.5	11.7	2247	US-60-173-464-13121	Sequence 13121, A
20	98.5	11.7	2285	US-60-191-637-16007	Sequence 16007, A
21	98.5	11.7	2285	US-60-191-661-12683	Sequence 12683, A
22	93	11.0	863	US-09-252-991A-21831	Sequence 21831, A
23	92.5	10.9	185	US-60-196-718-6866	Sequence 6866, Ap
24	91.5	10.8	275	US-09-733-089-18395	Sequence 18395, A
25	91	10.8	333	US-09-270-767-38937	Sequence 38937, A
26	91	10.8	333	US-09-270-767-54154	Sequence 54154, A
27	90.5	10.7	800	US-60-167-217-18636	Sequence 18636, A
28	90.5	10.7	800	US-60-173-464-15281	Sequence 15281, A
29	90.5	10.7	816	US-60-191-637-18596	Sequence 18596, A
30	90.5	10.7	816	US-60-191-661-14717	Sequence 14717, A
31	90	10.7	506	US-09-619-049-879	Sequence 879, App
32	90	10.7	506	US-60-167-217-17261	Sequence 17261, A
33	90	10.7	506	US-60-171-627-1365	Sequence 1365, Ap
34	90	10.7	506	US-60-173-464-14141	Sequence 14141, A
35	90	10.7	506	US-60-191-637-17261	Sequence 17261, A
36	90	10.7	506	US-60-191-661-13664	Sequence 13664, A
37	90	10.7	1572	US-60-224-729-2	Sequence 2, Appli
38	89	10.5	131	US-60-197-873-13405	Sequence 13405, A
39	89	10.5	259	US-60-167-217-10712	Sequence 10712, A
40	89	10.5	259	US-60-173-464-8662	Sequence 8662, Ap
41	89	10.5	259	US-60-191-637-10718	Sequence 10718, A
42	89	10.5	259	US-60-191-661-8396	Sequence 8396, Ap
43	88.5	10.5	415	US-09-421-239-8	Sequence 8, Appli
44	88.5	10.5	131	US-60-197-873-13419	Sequence 13419, A
45	88.5	10.5	152	US-60-197-873-13417	Sequence 13417, A

ALIGNMENTS

RESULT 1

US-09-125-635-8

Sequence 8, Application US/09125635

GENERAL INFORMATION:

APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE

TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator

FILE REFERENCE: 49944

CURRENT APPLICATION NUMBER: US/09/125, 635

PRIOR FILING DATE: 1998-08-21

PRIOR APPLICATION NUMBER: 60/049, 728

PRIOR FILING DATE: 1997-06-17

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 8

LENGTH: 951

TYPE: PRT

ORGANISM: Homo sapiens

US-09-125-635-8

Query Match	Score	DB 15:	Length	951:
Best Local Similarity	100.0%	Pred. No. 1.5e-79;		
Matches 162; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	EGVSHGTQNRPLRLNSLDDLVCPPNMLEGSDRALLDQTLTSTWDTGTEIDRALG	60	
DB	719	EGVSHGTQNRPLRLNSLDDLVCPPNMLEGSDRALLDQTLTSTWDTGTEIDRALG	778	
QY	61	IPELVNGGALFEBKQAFQGEAAVMDQKAGLYGCTGYAQCGRPMGCGFNLGQSPSPNS	120	
DB	779	IPELVNGGALFEBKQAFQGEAAVMDQKAGLYGCTGYAQCGRPMGCGFNLGQSPSPNS	838	
QY	121	MMQMMQCGFPLDQGMHPRANIMRPRTNPKQLRMQLQRLQ	162	

Db 839 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 880

RESULT 2

US-09-041-994-2

Sequence 2, Application US/09041994

GENERAL INFORMATION:

APPLICANT: Chen, J. Don

APPLICANT: Li, Hui

TITLE OF INVENTION: Transcriptional Coactivator for Nuclear

TITLE OF INVENTION: Hormone Receptors

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lanhive and Cockfield

STREET: 28 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/041,994

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Liepmann, W. Hugo

REGISTRATION NUMBER: 20,407

REFERENCE/DOCKET NUMBER: UMM-026-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-227-7400

TELEFAX: 617-742-4214

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1415 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-041-994-2

Query Match 100.0%; Score 845; DB 14; Length 1415;

Best Local Similarity 100.0%; Pred. No. 2,7e-79;

Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGVSHGTQNRPLRNSLDDLVGPPSNLEGQSDERALLDQLHTLLSNTDATGLEETDRALG 60

Db 1018 EGVSHGTQNRPLRNSLDDLVGPPSNLEGQSDERALLDQLHTLLSNTDATGLEETDRALG 1077

QY 61 IPELVNQGQALEPKQDAFGQGEAAVMMDQKAGLYGQTPAGQPPWQGFHLOGQSPSPNS 120

Db 1078 IPELVNQGQALEPKQDAFGQGEAAVMMDQKAGLYGQTPAGQPPWQGFHLOGQSPSPNS 1137

QY 121 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 162

Db 1138 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 1179

RESULT 3

US-09-513-066-2

Sequence 2, Application US/09513066

GENERAL INFORMATION:

APPLICANT: Chen, J. Don

APPLICANT: Leo, Christopher

APPLICANT: Li, Hui

TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF

TITLE OF INVENTION: STEROID NUCLEAR RECEPTORS

FILE REFERENCE: UMG-026CP

CURRENT APPLICATION NUMBER: US/09/513,066

CURRENT FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: USSN 09/041,994

PRIOR FILING DATE: 1998-03-13

PRIOR APPLICATION NUMBER: USSN 60/073,674

PRIOR FILING DATE: 1998-02-04

NUMBER OF SEQ ID NOS: 37

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 1415

TYPE: PRT

ORGANISM: Homo sapiens

US-09-513-066-2

Query Match 100.0%; Score 845; DB 19; Length 1415;

Best Local Similarity 100.0%; Pred. No. 2,7e-79;

Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGVSHGTQNRPLRNSLDDLVGPPSNLEGQSDERALLDQLHTLLSNTDATGLEETDRALG 60

Db 1018 EGVSHGTQNRPLRNSLDDLVGPPSNLEGQSDERALLDQLHTLLSNTDATGLEETDRALG 1077

QY 61 IPELVNQGQALEPKQDAFGQGEAAVMMDQKAGLYGQTPAGQPPWQGFHLOGQSPSPNS 120

Db 1078 IPELVNQGQALEPKQDAFGQGEAAVMMDQKAGLYGQTPAGQPPWQGFHLOGQSPSPNS 1137

QY 121 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 162

Db 1138 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 1179

RESULT 4

US-09-125-635-4

Sequence 4, Application US/09125635

GENERAL INFORMATION:

APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE

TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator

FILE REFERENCE: 49944

CURRENT FILING DATE: US/09/125,635

PRIOR APPLICATION NUMBER: 1998-08-21

PRIOR FILING DATE: 1997-06-17

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 1420

TYPE: PRT

ORGANISM: Homo sapiens

US-09-125-635-4

Query Match 100.0%; Score 845; DB 15; Length 1420;

Best Local Similarity 100.0%; Pred. No. 2,7e-79;

Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGVSHGTQNRPLRNSLDDLVGPPSNLEGQSDERALLDQLHTLLSNTDATGLEETDRALG 60

Db 1018 EGVSHGTQNRPLRNSLDDLVGPPSNLEGQSDERALLDQLHTLLSNTDATGLEETDRALG 1077

QY 61 IPELVNQGQALEPKQDAFGQGEAAVMMDQKAGLYGQTPAGQPPWQGFHLOGQSPSPNS 120

Db 1078 IPELVNQGQALEPKQDAFGQGEAAVMMDQKAGLYGQTPAGQPPWQGFHLOGQSPSPNS 1137

QY 121 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 162

Db 1138 MNQMNQGNFPLQGMHPRANIMRPTNPKQLRMQLQORLQ 1179

RESULT 5

US-09-125-635-12

Sequence 12, Application US/09125635

GENERAL INFORMATION:

APPLICANT: THE UNITED STATES OF AMERICA represented by THE SE

```
; TITLE OF INVENTION: AIB1, A novel steroid receptor co-activator
; FILE REFERENCE: 49944
; CURRENT APPLICATION NUMBER: US/09/125,635
; CURRENT FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 60/049,728
; PRIOR FILING DATE: 1997-06-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-125-635-12
```

```
Query Match          90.1%; Score 761; DB 15; Length 1402;
Best Local Similarity 89.5%; Pred. No. 1.7e-70;
Matches 145; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 EGVSHGTQNRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALG 60
DB 1029 EGGPHGSQNRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALG 1088
QY 61 IPELVNQGALPEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMOGFGHLOGSPSPFNS 120
DB 1089 IPELVNQGALPEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMOGFGHLOGSPSPFNS 1148
QY 121 MMNQMNQGNFPLQGMHPRANIMRPTNTPKQLRMOLQORLQ 162
DB 1149 MMQOISQOGSFPLOGMHPRAGLVPRPTNTPKQLRMOLQORLQ 1190
```

```
RESULT 6
US-09-513-066-14
; Sequence 14, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Leo, Christopher
; APPLICANT: Li, Hui
; TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
; FILE REFERENCE: UMG-026CP
; CURRENT APPLICATION NUMBER: US/09/513,066
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: USSN 09/041,994
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: USSN 60/073,674
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-066-14
```

```
Query Match          89.6%; Score 757; DB 19; Length 1398;
Best Local Similarity 88.9%; Pred. No. 4.6e-70;
Matches 144; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 1 EGVSHGTQNRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALG 60
DB 1025 EGGPHGSQNRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALG 1084
QY 61 IPELVNQGALPEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMOGFGHLOGSPSPFNS 120
DB 1085 IPELVNQGALPEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMOGFGHLOGSPSPFNS 1144
QY 121 MMNQMNQGNFPLQGMHPRANIMRPTNTPKQLRMOLQORLQ 162
DB 1145 MMQOISQOGSFPLOGMHPRAGLVPRPTNTPKQLRMOLQORLQ 1186
```

```
RESULT 7
US-09-445-353A-2
; Sequence 2, Application US/09445353A
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, et al., Michael
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
; FILE REFERENCE: 6627-PA1021
; CURRENT APPLICATION NUMBER: US/09/445,353A
; CURRENT FILING DATE: 1998-06-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-445-353A-2
```

```
Query Match          89.6%; Score 757; DB 18; Length 1402;
Best Local Similarity 88.9%; Pred. No. 4.6e-70;
Matches 144; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 1 EGVSHGTQNRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALG 60
DB 1025 EGGPHGSQNRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALG 1084
QY 61 IPELVNQGALPEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMOGFGHLOGSPSPFNS 120
DB 1085 IPELVNQGALPEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMOGFGHLOGSPSPFNS 1144
QY 121 MMNQMNQGNFPLQGMHPRANIMRPTNTPKQLRMOLQORLQ 162
DB 1145 MMQOISQOGSFPLOGMHPRAGLVPRPTNTPKQLRMOLQORLQ 1186
```

```
RESULT 8
US-09-513-066-15
; Sequence 15, Application US/09513066
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; APPLICANT: Leo, Christopher
; APPLICANT: Li, Hui
; TITLE OF INVENTION: NOVEL USES FOR THE RAC3 TRANSCRIPTIONAL COACTIVATOR OF
; FILE REFERENCE: UMG-026CP
; CURRENT APPLICATION NUMBER: US/09/513,066
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: USSN 09/041,994
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: USSN 60/073,674
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-066-15
```

```
Query Match          66.8%; Score 564.5; DB 19; Length 1391;
Best Local Similarity 71.9%; Pred. No. 7.2e-50;
Matches 115; Conservative 17; Mismatches 25; Indels 3; Gaps 3;
```

```
QY 6 GTONRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALGIPELY 65
DB 1014 GAONRPLRLNSLDLVPSPNLEGGSDERALLDQHTLLSNTDGTGLEIDRALGIPELY 1073
QY 66 NOGQALEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMO-GGF-HLOGQSPSPFNSMN 123
DB 1074 SOGQALEPKODAFQGEAAVMMQKAGLYGQTYPAQGPMPMO-GGF-HLOGQSPSPFNSMN 1133
```

QY 124 OMN-OQGNFPIQGMHPRANIMPRNTKQLRMQIQRLQ 162
 DB 1134 OMNOQGMHPIQGMHPRANILIRNNIKQLRMQIQRLQ 1173

RESULT 9

US-60-258-273-113
 ; Sequence 113, Application US/60258273
 ; GENERAL INFORMATION:
 ; APPLICANT: Beasley, Ellen
 ; TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
 ; TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
 ; FILE REFERENCE: C1001042-PROV
 ; CURRENT APPLICATION NUMBER: US/60/258, 273
 ; CURRENT FILING DATE: 2000-12-27
 ; NUMBER OF SEQ ID NOS: 312
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 113
 ; LENGTH: 1127
 ; TYPE: PRT
 ; ORGANISM: Human
 US-60-258-273-113

Query Match 37.0%, Score 312.5; DB 23; Length 1127;
 Best Local Similarity 44.1%, Pred. No. 1.5e-23;
 Matches 75; Conservative 23; Mismatches 49; Indels 23; Gaps 5;

QY 1 ECVSHGTGNRPLRNSLDDLVGPSPNLEGSODERALLDQLHTLSNTDPTGLEIDRALG 60
 DB 692 DQASFASQNRQPFSSPDLLCPHRAESPSDEGALLDQLVTLALNFD--GLEIDRALG 749
 QY 61 IPELVNOCALPEPKODAFQGEAAVMMDOKAGLYGQTYPAQGPWMQ-----GFHLQ 112
 DB 750 IPELVQSQAQVADPEQ--FSSQDSNIMLEQKAPVFPOQYASQAQMAQGSYSPMDPNFHTM 807
 QY 113 GQSPFSNMMNOMNOQGNFPIQGMHPRANIMPRNTKQLRMQIQRLQ 162
 DB 808 GQSPFSATLRMQ-----PRGLRPTGLVQ---NQPQRLQLDQHLRLQ 846

RESULT 10

US-08-891-640-2
 ; Sequence 2, Application US/08891640
 ; GENERAL INFORMATION:
 ; APPLICANT: Chambon, Pierre
 ; APPLICANT: Gronemeyer, Hinrich
 ; APPLICANT: Voegel, Johannes
 ; APPLICANT: Lutz, Yves
 ; TITLE OF INVENTION: Transcriptional Intermediary Factor-2
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: Steiner, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, NW, Suite 600
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3934
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/891,640
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/021,247
 ; FILING DATE: 12-JUL-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Steffe, Eric K.

REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1383.0130001/ERS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1464 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-891-640-2

Query Match 37.0%, Score 312.5; DB 12; Length 1464;
 Best Local Similarity 44.1%, Pred. No. 2.2e-23;
 Matches 75; Conservative 23; Mismatches 49; Indels 23; Gaps 5;

QY 1 ECVSHGTGNRPLRNSLDDLVGPSPNLEGSODERALLDQLHTLSNTDPTGLEIDRALG 60
 DB 1044 DQASFASQNRQPFSSPDLLCPHRAESPSDEGALLDQLVTLALNFD--GLEIDRALG 1101
 QY 61 IPELVNOCALPEPKODAFQGEAAVMMDOKAGLYGQTYPAQGPWMQ-----GFHLQ 112
 DB 1102 IPELVQSQAQVADPEQ--FSSQDSNIMLEQKAPVFPOQYASQAQMAQGSYSPMDPNFHTM 1159
 QY 113 GQSPFSNMMNOMNOQGNFPIQGMHPRANIMPRNTKQLRMQIQRLQ 162
 DB 1160 GQSPFSATLRMQ-----PRGLRPTGLVQ---NQPQRLQLDQHLRLQ 1198

RESULT 11

US-09-445-353A-3
 ; Sequence 3, Application US/09445353A
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosenfeld, et al., Michael
 ; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
 ; FILE REFERENCE: 6627-PA1021
 ; CURRENT APPLICATION NUMBER: US/09/445,353A
 ; CURRENT FILING DATE: 1998-06-12
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 1462
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-445-353A-3

Query Match 36.7%, Score 310.5; DB 18; Length 1462;
 Best Local Similarity 46.1%, Pred. No. 3.6e-23;
 Matches 76; Conservative 26; Mismatches 50; Indels 13; Gaps 6;

QY 1 ECVSHGTGNRPLRNSLDDLVGPSPNLEGSODERALLDQLHTLSNTDPTGLEIDRALG 60
 DB 1044 DQASFASQNRQPFSSPDLLCPHRAESPSDEGALLDQLVTLALNFD--GLEIDRALG 1101
 QY 61 IPELVNOCALPEPKODAFQGEAAVMMDOKAGLYGQTYPAQGPWMQGFHLQSGSPFSNS 120
 DB 1102 IPELVQSQAQVADPEQ--FSSQDSNIMLEQKAPVFPOQYASQAQMAQGSYSPMDPNFHTM 1158
 QY 121 MNOMNOQGNFPIQGMHPRANIMPR---TTPKQLRMQIQRLQ 162
 DB 1159 ---WGQRPNTYTLRMQRRP--LRPTGLVQNOPQLRLDQHLRLQ 1198

RESULT 12

US-08-891-640-9
 ; Sequence 9, Application US/08891640
 ; GENERAL INFORMATION:
 ; APPLICANT: Chambon, Pierre
 ; APPLICANT: Gronemeyer, Hinrich
 ; APPLICANT: Voegel, Johannes

APPLICANT: Lutz, Yves
TITLE OF INVENTION: Transcriptional Intermediary Factor-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,640
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/021,247
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1383.0130001/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-891-640-9

Query Match 32.9%; Score 278; DB 12; Length 58;
Best Local Similarity 94.8%; Pred. No. 8.3e-22;
Matches 55; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 LKNSLDDLVGPPSNLEGSODERALLDQLHTLSNTDATGTELRALGIPELVNGOQA 70
DB 1 LKNSLDDLVGPPSNLEGSODERALLDQLHTLSNTDATGTELRALGIPELVNGOQA 58

RESULT 13
US-08-891-640-3
Sequence 3, Application US/08891640
GENERAL INFORMATION:
APPLICANT: Chambon, Pierre
APPLICANT: Gronemeyer, Hinrich
APPLICANT: Voegel, Johannes
APPLICANT: Lutz, Yves
TITLE OF INVENTION: Transcriptional Intermediary Factor-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/891,640
FILING DATE: Herewith

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/021,247
FILING DATE: 12-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1383.0130001/EKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1036 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-891-640-3

Query Match 31.5%; Score 266.5; DB 12; Length 1036;
Best Local Similarity 42.5%; Pred. No. 8.9e-19;
Matches 68; Conservative 28; Mismatches 49; Indels 15; Gaps 7;

QY 13 LKNSLDDLVGPPSNLEGSODERALLDQLHTLSNTDATGTELRALGIPELVNGOQA 72
DB 535 ISSQLDELCPPTVEGRNDEKALLEQLVPSGKDETELAEDRLGIDKLIV-QGGGLD 593

QY 73 PKDAFQGOEAA-VAMDQKAGLYGQTYPAQGP-----PMQGGFHLQGSQSFNSMMNQ 124
DB 594 VLSERPPOGATPPLIMERPNLYSQPYSSPFTANLSPFPG---MVKQKSLGTMPQ 650

QY 125 MN-QQGNF-PLOGMHPRANIMPRNTPKOLRMQLQORIQ 162
DB 651 VTPRGAFSPGMQMPRQTLNRPPA-APNQLRLQLOQRQ 689

RESULT 14
US-08-701-154A-5
Sequence 5, Application US/08701154A
GENERAL INFORMATION:
APPLICANT: Tsai, Ming-Jer
APPLICANT: Tsai, Sophia Y.
APPLICANT: Onate, Sergio A.
TITLE OF INVENTION: STEROID RECEPTOR COACTIVATOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,154A
FILING DATE: August 21, 1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/003,784
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 220/243
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1061 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-701-154A-5

Query Match

31.5%; Score 266.5; DB 11; Length 1061;

Best Local Similarity 42.5%; Pred. No. 9.3e-19; Mismatches 49; Indels 15; Gaps 7;

Matches 68; Conservative 28; Mismatches 49; Indels 15; Gaps 7;

QY 13 LRNSLDLVGPPSNLEGSDERALDQLHTLLSNTDATGLEIDRALGIPELVNOGALE 72

Db 529 ISSOLDELICPPTVEGRNDEKALLEQLVSLSGKDETELAEIDRALGIDRLV-QGGGLD 587

QY 73 PKDAFQGOEAA--VMMDQKAGLYGQRTPAQGP-----PMOGGFHLQGSFNSMMNQ 124

Db 588 VLSEFPPOQATPPLIMEERNLYSQPYSSPPTANLPSPPG--WVRQKPSLGTMPVQ 644

QY 125 MN-QQGNF-PLQGMHPRANIMRPTNTPKQLRMQLQQRLO 162

Db 645 VTPRGAFSPGMQPRQTLNRPPA-APNQLRLQLQQRLO 683

RESULT 15

US-60-160-189-8436

; Sequence 8436, Application US/60160189

; GENERAL INFORMATION:

; APPLICANT: BONAZZI, VIVIEN

; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CLO00112

; CURRENT APPLICATION NUMBER: US/60/160,189

; CURRENT FILING DATE: 1999-10-19

; NUMBER OF SEQ. ID NOS: 10162

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ. ID NO 8436

; LENGTH: 51

; TYPE: PRT

; ORGANISM: HUMAN

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)..(51)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-60-160-189-8436

Query Match

28.0%; Score 237; DB 23; Length 51;

Best Local Similarity 92.2%; Pred. No. 1.4e-17;

Matches 47; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 NRPLLRNSLDLVGPPSNLEGSDERALDQLHTLLSNTDATGLEIDRAL 59

Db 1 SRPLLRNSLDLVGPPSTLEGSDERALLDQXHTLVSNMTDATGLEIDRAL 51

Search completed: April 19, 2001, 16:54:06
Job time: 493 sec

FILE REFERENCE: 38-10(15490)C
CURRENT APPLICATION NUMBER: US/09/739,449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 8151
LENGTH: 407
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-8151

Query Match 7.4%; Score 62.5; DB 5; Length 407;
Best Local Similarity 23.5%; Pred. No. 11;
Matches 39; Conservative 20; Mismatches 58; Indels 49; Gaps 8;

QY 12 LIRNSLDLVGPPSNLEGSQSDERALLDQLHTLLSNTDATGTEIDRALGPELVNOGO-A 70
DB 44 ILRGALDGLTAVPP--VGDPEYALRQOL-FLRTEGEN-----PALALDGYFA 88
QY 71 LEPR-----QDAFGGGEAAVMDQKAGLYGQTPPAOGPPMOGFGHLOG--QSFNSMNMNO 124
DB 89 LHPSLRNFORLYRKGEAAVNHASATGYDRSHFDGODYLESGFATPGHESGMLNRLLEQ 148
QY 125 MNO-----OGNFPLOGMHPRANIMRP 145
DB 149 MPAGETIPNAGCDRVSGIATGASAPLVIRGKAPILGMAP--SVLRP 192

RESULT 7
US-09-661-322A-28
Sequence 28, Application US/09661322A
GENERAL INFORMATION:
APPLICANT: Baum, James A.
APPLICANT: Chu, Chih-Rel
APPLICANT: Donovan, William P.
APPLICANT: Gilmer, Amy J.
APPLICANT: Ruper, Mark J.
TITLE OF INVENTION: Leptodipteran-Active Bacillus thuringiensis Delta-Endotoxin Compo
FILE OF INVENTION: and Methods of Use
FILE REFERENCE: MECO201
CURRENT APPLICATION NUMBER: US/09/661,322A
CURRENT FILING DATE: 2000-09-13
NUMBER OF SEQ ID NOS: 63
SOFTWARE: PatentIn version 3.0
SEQ ID NO 28
LENGTH: 1156
TYPE: PRT
ORGANISM: Bacillus thuringiensis
US-09-661-322A-28

Query Match 7.4%; Score 62.5; DB 5; Length 1156;
Best Local Similarity 26.0%; Pred. No. 43;
Matches 39; Conservative 18; Mismatches 48; Indels 45; Gaps 9;

QY 21 VPPSNLEGSQSDERALLD--QLHTLLSNTDATGTEIDRA-LGPELVN-----OGAL 71
DB 912 VGP--LGESEIERORONAKMNMELGRKRA---EIDRVYLAQAQAINHLEFVDYQDOL 964
QY 72 EPRKDAFGGGEAAVMDQKAGLYGOT-----YPAOGPPMOGFGHLOG-----OS 115
DB 965 NPEIGLAEINERNSLVESISGYSTLLQIPGINTETTESDRLOQASLYLTSNNAYON 1024
QY 116 PSFNSMNMNM-----OQGNFPLOGMH 137
DB 1025 GDFNSGLDSMNTTMDASVQODCN-----MH 1049

RESULT 8
US-09-739-449-10802
Sequence 10802, Application US/09739449

GENERAL INFORMATION:
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15490)C
CURRENT APPLICATION NUMBER: US/09/739,449
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 09/514,000
PRIOR FILING DATE: 2000-02-23
NUMBER OF SEQ ID NOS: 13351
SEQ ID NO 10802
LENGTH: 181
TYPE: PRT
ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10802

Query Match 7.3%; Score 62; DB 5; Length 181;
Best Local Similarity 32.6%; Pred. No. 4.3;
Matches 28; Conservative 12; Mismatches 30; Indels 16; Gaps 5;

QY 14 RNSLDLVGPPSNLEGSQSDERALLD--QLHTLL--SNTDATGTEIDRALGPELVNO 67
DB 31 RQCLTEALG---NLGEPDSLLTDMQSRSLDLWKHSNGSARSISIKWIDRA---AVLICR 84
QY 68 GQALEPKDAPGQGEAAVMDQKAGL 93
DB 85 DYLLENSADTVVSOE---LEQISGL 106

RESULT 9
US-09-819-142-24
Sequence 24, Application US/09819142
GENERAL INFORMATION:
APPLICANT: Reuber, Lynne
APPLICANT: Ratcliffe, Oliver
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Riechmann, Jose Luis
APPLICANT: Heard, Jacqueline
TITLE OF INVENTION: Methods for Modifying Flowering Phenotypes
FILE OF INVENTION: MBI-0033
FILE REFERENCE: MBI-0033
CURRENT APPLICATION NUMBER: US/09/819,142
CURRENT FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24
LENGTH: 669
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-819-142-24

Query Match 7.3%; Score 62; DB 5; Length 669;
Best Local Similarity 24.4%; Pred. No. 24;
Matches 49; Conservative 22; Mismatches 70; Indels 60; Gaps 13;

QY 7 TQNRPLRNSL---DDLVPSPSNLEGSQSDERALLDQLHTLLSNTDATGTEIDRA-LGI 61
DB 299 SOHQGNLNNSEFTGQDAFGLSTLNG-FDLQAL-----AVTGQLPQASLQLOQAAGGR 352
QY 62 PELVNOG---QALEPKDAPGQGEAAVMDQKAGLYGQTPPAOGPPMOGFGH-----LQ 112
DB 353 PAMWSKGLPVSSIVDERSTFSFNTKTRFEGELGHQO-QPQOQPM-NLHGVPTGLQ 410
QY 113 GQSPFSNMMNMNM-----OQGN-----FPLOGMH--PR-----ANIMRP--- 145
DB 411 QQLPWGNRMSIQOQIAAAYRAGNSVYNNGLMPLAGQOSLPAGPPMLTSSQSSIRPMLS 470
QY 146 -----RTNTPKQLRM 155
DB 471 NRISERSGSGRRNNIPESSRV 491

DB 517 KSKKPGSVKTERAGAGQ-IEKOMVVRGGRVGRGRC-VNLAVLKLSBOG 565

RESULT 14

US-09-739-449-13145
; Sequence 13145, Application US/09739449
; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15480)C

; CURRENT APPLICATION NUMBER: US/09/739,449

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/514,000

; PRIOR FILING DATE: 2000-02-23

; NUMBER OF SEQ ID NOS: 13351

; SEQ ID NO 13145

; LENGTH: 365

; TYPE: PRT

; ORGANISM: Agrobacterium tumefaciens

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(365)

; OTHER INFORMATION: unsure at all Xaa locations

US-09-739-449-13145

Query Match 7.0%; Score 59; DB 5; Length 365;

Best Local Similarity 20.6%; Pred. No. 21;

Matches 34; Conservative 24; Mismatches 49; Indels 58; Gaps 8;

OY 19 DIVGPPSNLEGO-----SDERALDOLHTLSN--TDATGLEIDRALGPE 63

DB 133 DYKAPISGIVGARITGALVNTSGENLTIQOLDPIADFTQPRADIRLRLKAL---- 188

OY 64 LVNQGQALEPKODAFQGEAAVMDQKAGLYGQTPAGCPMGQGFHLQGSFSPNSMNN 123

DB 189 --QDCQLMTGQNEA---EVNLLFDD---GSRYPVSG-----RLFSEAAVDETTG 230

OY 124 QMNQGNFP-----LQGHPRANIMRPRNTTPKOLRMOLQRLQ 162

DB 231 QVTLRGEFPNPGDILPGMY-----VRVQIOOGIQ 260

RESULT 15

US-09-739-449-11226
; Sequence 11226, Application US/09739449
; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15480)C

; CURRENT APPLICATION NUMBER: US/09/739,449

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/514,000

; PRIOR FILING DATE: 2000-02-23

; NUMBER OF SEQ ID NOS: 13351

; SEQ ID NO 11226

; LENGTH: 306

; TYPE: PRT

; ORGANISM: Agrobacterium tumefaciens

US-09-739-449-11226

Query Match 6.9%; Score 58.5; DB 5; Length 306;

Best Local Similarity 23.9%; Pred. No. 19;

Matches 28; Conservative 12; Mismatches 34; Indels 43; Gaps 5;

OY 23 PSNLEG--QSDERALDOLHTLSN-----DATGLEIDR--ALGIPELVNOGAL 71

DB 96 PLSNIDALRALDMRIELTNLHRTKATMTYVTHDOVEAMTMADRIYVNLNAGEIAOVGAPL 155

OY 72 E-----PKOD-----AFQGEAAVMDQKAGLYGQ 96

DB 156 ELYHKRANLFVAGFIGNPKMFLKVTCKSVSAEGVTVAVEGQITIVPVEPRAGLEBK 212

Search completed: April 19, 2001, 16:54:13
Job time: 497 sec

